

PCT

世界知的所有権機関
国際事務局

特許協力条約に基づいて公開された国際出願



B69

(51) 国際特許分類7 C12N 15/55, 9/16, C07K 16/40, C12Q 1/68, A61K 38/46	A1	(11) 国際公開番号 WO00/63392 (43) 国際公開日 2000年10月26日(26.10.00)
(21) 国際出願番号 PCT/JP00/02455 (22) 国際出願日 2000年4月14日(14.04.00) (30) 優先権データ 特願平11/108842 1999年4月16日(16.04.99) JP (71) 出願人 (米国を除くすべての指定国について) 協和醗酵工業株式会社 (KYOWA HAKKO KOGAKU CO., LTD.)[JP/JP] 〒100-8185 東京都千代田区大手町一丁目6番1号 Tokyo, (JP) (72) 発明者 ; および (75) 発明者 / 出願人 (米国についてのみ) 清水憲二(SHIMIZU, Kenji)[JP/JP] 〒703-8282 岡山県岡山市平井三丁目592番6号 Okayama, (JP)	(81) 指定国 AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, 欧州特許 (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI特許 (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG), ARIPO特許 (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), ユーラシア特許 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM) 添付公開書類 国際調査報告書 明細書とは別に規則13の2に基づいて提出された生物材料の寄託に関する表示。	
(54) Title: NOVEL TYROSINE PHOSPHATASE (54) 発明の名称 新規チロシンフォスファターゼ (57) Abstract A novel tyrosine phosphatase encoded by the human chromosome 3p21-site the deletion of which is frequently observed in human lung cancer and thus which is assumed as having an antioncogen therein. Use of this tyrosine phosphatase and its gene makes it possible to diagnose and treat cancer.		

(57)要約

本発明は、ヒト肺癌において高頻度に欠失が観察され、癌抑制遺伝子の存在が推定されているヒト染色体3p21部位にコードされる、新規チロシンフォスファターゼに関する。該チロシンフォスファターゼおよびその遺伝子を用いることにより、癌の診断および治療が可能となる。

PCTに基づいて公開される国際出願のパンフレット第一頁に掲載されたPCT加盟国を同定するために使用されるコード(参考情報)

AE	アラブ首長国連邦	DM	ドミニカ	KZ	カザフスタン	RU	ロシア
AG	アンティグア・バーブーダ	DZ	アルジェリア	LC	セントルシア	SD	スーダン
AL	アルバニア	EE	エストニア	LI	リヒテンシュタイン	SE	スウェーデン
AM	アルメニア	ES	スペイン	LK	スリ・ランカ	SG	シンガポール
AT	オーストリア	FI	フィンランド	LR	リベリア	SI	スロヴェニア
AU	オーストラリア	FR	フランス	LS	レソト	SK	スロヴァキア
AZ	アゼルバイジャン	GA	ガボン	LT	リトアニア	SL	シエラ・レオネ
BA	ボスニア・ヘルツェゴビナ	GB	英国	LU	ルクセンブルグ	SN	セネガル
BB	バルバドス	GD	グレナダ	LV	ラトヴィア	SZ	スワジランド
BE	ベルギー	GE	グルジア	MA	モロッコ	TD	チャード
BF	ブルキナ・ファソ	GH	ガーナ	MC	モナコ	TG	トーゴ
BG	ブルガリア	GM	ガンビア	MD	モルドヴァ	TJ	タジキスタン
BJ	ベナン	GN	ギニア	MG	マダガスカル	TM	トルクメニスタン
BR	ブラジル	GR	ギリシャ	MK	マケドニア旧ユーゴスラヴィア共和国	TR	トルコ
BY	ベラルーシ	GW	ギニア・ビサウ	ML	マリ	TT	トリニダード・トバゴ
CA	カナダ	HR	クロアチア	MN	モンゴル	TZ	タンザニア
CF	中央アフリカ	HU	ハンガリー	MR	モーリタニア	UA	ウクライナ
CG	コンゴ	ID	インドネシア	MW	マラウイ	UG	ウガンダ
CH	スイス	IE	アイルランド	MX	メキシコ	US	米国
CI	コートジボアール	IL	イスラエル	MZ	モザンビーク	UZ	ウズベキスタン
CM	カメルーン	IN	インド	NE	ニジェール	VN	ヴェトナム
CN	中国	IS	アイスランド	NL	オランダ	YU	ユーゴスラヴィア
CR	コスタ・リカ	IT	イタリア	NO	ノルウェー	ZA	南アフリカ共和国
CU	キューバ	JP	日本	NZ	ニュー・ジーランド	ZW	ジンバブエ
CY	キプロス	KE	ケニア	PL	ポーランド		
CZ	チェコ	KG	キルギスタン	PT	ポルトガル		
DE	ドイツ	KP	北朝鮮	RO	ルーマニア		
DK	デンマーク	KR	韓国				

明 細 書

新規チロシンフォスファターゼ

技術分野

本発明は癌抑制遺伝子と推定される、ヒト染色体3p21にコードされている新規チロシンフォスファターゼに関する。

背景技術

癌抑制遺伝子の不活性化は、癌の発症に重要な役割を果たしている。ゲノム上の遺伝子是对立遺伝子として2コピー存在するので、癌抑制遺伝子の不活性化は、多くの場合、一方の遺伝子に染色体欠失 (loss of heterozygosity : 以下、LOHと略す)、もう一方の遺伝子に突然変異がおこることによるものとされている。したがって、癌細胞において高頻度にLOHの観察される部位には癌抑制遺伝子が存在している可能性が高いと考えられる。ヒトの肺癌において、LOHが高頻度でおこる部位が報告されているが、その中でも第3染色体短腕 (3p) は、肺の小細胞癌の100%近くに、また非小細胞癌においても60%に欠失が見られることから、肺癌の発症に関与する遺伝子の存在部位と考えられている〔Oncogene, 7, 445 (1992)〕。LOHは3pのいろいろな領域で観察されているが、その中でも3p14、3p21、3p25は多くの肺癌で共通して欠失が見られる領域であり、これらの領域には癌抑制遺伝子が存在すると推定されている〔Oncogene, 7, 445 (1992)〕。これらの領域の欠失は肺癌発症過程の初期からみられるが、悪性化に伴い欠失範囲が拡大していくという報告〔Oncogene, 11, 2591 (1995)〕や、非小細胞癌、特に肺腺癌の不良な予後因子である報告もある。これまでにこれらの領域にマッピングされた癌抑制遺伝子として、3p14ではFHIT (Fragile histidine triad) 遺伝子〔Cell, 84, 587 (1996)、Cell, 85, 17 (1996)〕が、3p25ではVHL (von Hippel-Lindau) 遺伝子〔Science, 260, 1317 (1993)、Proc. Natl. Acad.

Sci., 88, 2864 (1991)) が報告されている。3p21では、DNAミスマッチ修復遺伝子MLH1 [MutL homologue; Nature, 368, 258 (1994)]、アミノアシラーゼ1 [Genomics, 8, 149 (1990)、J. Biol. Chem., 268, 1710 (1993)]、3pK [mitogen-activated protein kinase-activated protein kinase; Chromosome Research, 4, 310 (1996)]、セマフォリン [Oncogene, 12, 1289 (1996)]、アルギニンリッチ蛋白質 [Oncogene, 12, 1931 (1996)]、ユビキチン活性化蛋白質、Wnt-5a [J. Biol. Chem., 270, 31225 (1995)、Cell Growth & Differentiation, 8, 417 (1997)] などが候補遺伝子としてマッピングされているが、肺癌での遺伝子変異と関連付けられる決定的な癌抑制遺伝子は見出されていない。また、ヒト第3染色体とマウスA9細胞の雑種融合細胞クローンのうち、SCIDマウスに移植した場合に腫瘍形成を示すような2種類の雑種融合細胞クローンにおいて、共通して欠失しているヒト第3染色体上の領域を解析した論文 [Genes, Chromosome & Cancer, 18, 200 (1997)、Genes, Chromosome & Cancer, 20, 329 (1997)] では、欠失部分はヒト第3染色体上の3p21.3の領域内であり、領域の大きさとして1.6cM (センチモルガン) まで (染色体マーカーD3S1029とD3S643の間) 特定されている。この欠失領域には腫瘍形成に直接関与する癌抑制遺伝子が存在すると考えられているが、その遺伝子は特定されるまでには至っていない。

発明の開示

ヒトの肺癌において高頻度に欠失が観察される、ヒト染色体3p21部位にコードされている癌抑制遺伝子を明らかにすることが望まれている。

本発明者らは、高頻度に欠失が観察される、ヒト染色体3p21部位にコードされている癌抑制遺伝子に関して鋭意検討し、癌抑制に関与する新規チロシンフォスファターゼHD-PTP (histidine domain-protein tyrosine phosphatase; 以下、HD-PTPと略すこともある) をコードするcDNAをヒト細胞株よりクローン化

することに成功し、本発明を完成するに至った。

すなわち、本発明は、以下の(1)～(20)を提供するものである。

- (1) 配列番号2記載のアミノ酸配列からなる蛋白質。
- (2) 配列番号2記載の蛋白質の有するアミノ酸配列において1個以上のアミノ酸が欠失、置換もしくは付加されたアミノ酸配列からなり、かつチロシンフォスファターゼ活性を有する蛋白質。

上記(2)の配列番号2記載の蛋白質の有するアミノ酸配列において1個以上のアミノ酸が欠失、置換もしくは付加されたアミノ酸配列からなり、かつチロシンフォスファターゼ活性を有する蛋白質は、Molecular Cloning, A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press (1989) (以下、モレキュラー・クローニング第2版と略す)、Current Protocols in Molecular Biology, John Wiley & Sons (1987-1997) (以下、カレント・プロトコールズ・イン・モレキュラー・バイオロジーと略す)、Nucleic Acids Research, 10, 6487 (1982)、Proc. Natl. Acad. Sci. USA, 79, 6409(1982)、Gene, 34, 315 (1985)、Nucleic Acids Research, 13, 4431 (1985)、Proc. Natl. Acad. Sci. USA, 82, 488 (1985)等に記載の部位特異的変異導入法を用いて、例えば配列番号2で示されるアミノ酸配列を有するポリペプチドをコードするDNAに部位特異的変異を導入することにより、取得することができる。

欠失、置換もしくは付加されるアミノ酸の数は特に限定されないが、上記の部位特異的変異法等の周知の方法により欠失、置換もしくは付加できる程度の数であり、1個から数十個、好ましくは1～20個、より好ましくは1～10個、さらに好ましくは1～5個である。

また、本発明のポリペプチドがチロシンフォスファターゼ活性を有するためには、配列番号2記載のアミノ酸配列と、BLAST [J. Mol. Biol., 215, 403 (1990)] やFASTA [Methods in Enzymology, 183, 63-98 (1990)] 等を用いて計算したときに、少なくとも60%以上、通常は80%以上、特に95%

以上の相同性を有していることが好ましい。

具体的には配列番号2のアミノ酸配列のN末端1～3番目の3アミノ酸を欠失したアミノ酸配列を有する蛋白質、この3アミノ酸が欠失したアミノ酸配列のN末にT7やFlag等のtagペプチドの配列を付加したアミノ酸配列を有する蛋白質などをあげることができる。

(3) 上記(1)または(2)記載の蛋白質をコードするDNA。

(4) 配列番号1、3、40または41記載の塩基配列を有するDNA。

(5) 上記(3)または(4)記載のDNAとストリンジントな条件下でハイブリダイズし、かつチロシンフォスファターゼ活性を有する蛋白質をコードするDNA。

上記(5)の「ストリンジントな条件下でハイブリダイズし、かつチロシンフォスファターゼ活性を有する蛋白質をコードするDNA」とは、配列番号1、3、40または41で表される塩基配列を有するDNAをプローブとして、コロニー・ハイブリダイゼーション法、ブラーク・ハイブリダイゼーション法あるいはサザン・プロット・ハイブリダイゼーション法等を用いることにより得られるDNAを意味し、具体的には、コロニーあるいはブラーク由来のDNAを固定化したフィルターを用いて、0.7～1.0 mol/LのNaCl存在下、65℃でハイブリダイゼーションを行った後、0.1×～2×SSC(saline-sodium citrate)溶液〔1×SSC溶液(150 mmol/L NaCl、15 mmol/Lクエン酸ナトリウム)；n×はn倍濃度の溶液を示す。〕を用い、65℃条件下でフィルターを洗浄することにより同定できるDNAをあげることができる。

ハイブリダイゼーションは、モレキュラー・クローニング第2版、カレント・プロトコールズ・イン・モレキュラー・バイオロジー、DNA Cloning 1: Core Techniques, A Practical Approach, Second Edition, Oxford University Press (1995) (以下、DNAクローニング1と略記する)等の実験書に記載されている方法に準じて行うことができる。ハイブリダイズ可能なDNAとして具体的には、配

列番号 1、3、40 または 41 で表される塩基配列と少なくとも 80% 以上の相同性を有する DNA、好ましくは 95% 以上の相同性を有する DNA をあげることができる。

(6) 上記 (3) ~ (5) のいずれか 1 項に記載の DNA とベクター DNA を含有する組換え体 DNA。

(7) 上記 (6) 記載の組換え体 DNA を宿主細胞に導入して得られる形質転換体。

(8) 上記 (7) 記載の形質転換体を培地に培養し、培養物中に上記 (1) または (2) 記載の蛋白質を生成蓄積させ、該培養物から該蛋白質を採取することを特徴とする上記 (1) または (2) 記載の蛋白質の製造方法。

(9) 上記 (1) または (2) 記載の蛋白質を有効成分として含む、癌の治療薬。

(10) 上記 (3) ~ (5) のいずれか 1 項に記載の DNA を有効成分とする癌の治療薬。

(11) 上記 (3) ~ (5) のいずれか 1 項に記載の DNA を含有する、癌の遺伝子治療用ベクター。

(12) 上記 (3) ~ (5) のいずれか 1 項に記載の DNA の塩基配列のうち、連続した 10~60 残基の塩基配列を有するオリゴヌクレオチド、または該オリゴヌクレオチドと相補的な配列を有するオリゴヌクレオチド、およびこれらオリゴヌクレオチドの誘導体であるオリゴヌクレオチド誘導体。

(13) 配列番号 24、25、28 または 29 で表される塩基配列を有する上記 (12) 記載のオリゴヌクレオチド。

(14) 上記 (12) または (13) 記載のオリゴヌクレオチドを用いた、癌の診断方法。

(15) 上記 (12) または (13) 記載のオリゴヌクレオチドを含有する、癌の診断薬。

(16) 上記(1)または(2)記載の蛋白質を認識する抗体。

(17) 上記(16)記載の抗体を用いて、上記(1)または(2)記載の蛋白質を免疫学的に検出する方法。

(18) 上記(16)記載の抗体を用いて、上記(1)または(2)記載の蛋白質を免疫学的に定量する方法。

(19) 上記(16)記載の抗体を用いる、癌の判定方法。

(20) 上記(16)記載の抗体を有効成分とする、癌の診断薬。

以下、本発明を詳細に説明する。

[1]本発明のDNAの調製

本発明のHD-PTPをコードするcDNAは、以下のような形質転換アッセイ法によって得られる、癌関連遺伝子の候補となる遺伝子の断片をプローブとしてcDNAライブラリーをスクリーニングすることにより取得することができる。

以下、詳細に説明する。

(1) 癌関連遺伝子断片のスクリーニング

ヒト癌細胞は、無限増殖性の獲得(不死化)、細胞増殖時の接触阻止現象の喪失、足場非依存性の増殖、増殖における血清要求性の低下、細胞接着能の減弱、ヌードマウスに移植した場合の腫瘍形成能など、正常細胞とは異なる性質を有しているので、ヒト癌細胞の染色体DNA中には、上述の性質の原因となる変異があると考えられる。このような原因遺伝子をスクリーニングする方法の一つとして、形質転換アッセイ法があげられる。

形質転換アッセイ法は、正常細胞、または上述した癌細胞の形質のうち、無限増殖以外の癌形質(造腫瘍性、軟寒天内増殖能など)を有さない正常細胞と同様の性質を示す宿主細胞に、ヒト癌細胞の染色体DNA由来のDNA断片を導入して発現させ、正常な形質からヒト癌細胞の示す形質への変化を指標に細胞クローンを選択し、該細胞から形質転換の原因となった染色体DNA断片(原因遺伝子)を単離する方法である。

ヒト癌細胞としては、染色体DNAを抽出できるヒト癌細胞であればいかなる細胞でも用いることができる。本発明では、ヒト癌細胞としてヒトB細胞リンパ腫の細胞株であるKAL-1〔Cancer Res., 51, 5392 (1991)〕を用いている。

ヒト癌細胞から染色体DNAを単離する方法としては、プロテイナーゼK／フェノールクロロホルム抽出法（モレキュラー・クローニング第2版）等をあげることができる。

上記で得られた染色体DNA由来のDNA断片を宿主細胞に導入する際には、導入された細胞を選択するためにマーカーとなる薬剤耐性遺伝子を有するプラスミドを共導入することが望ましい。該プラスミドとしては、マーカーとなる薬剤耐性を有しているプラスミドであればいかなるプラスミドでも用いることができる。具体的には、ハイグロマイシン耐性遺伝子を有するpHyg〔Mol. Cel. Biol., 5, 410 (1985)〕などをあげることができる。

染色体DNA由来のDNA断片とプラスミドの導入方法としては、宿主細胞へDNAを導入する方法であればいずれも用いることができるが、例えば、エレクトロポレーション法〔Cytotechnology, 3, 133 (1990)〕、リン酸カルシウム法（特開平2-227075）、リポフェクション法〔Proc. Natl. Acad. Sci. USA, 84, 7413 (1987)〕等をあげることができる。

宿主細胞としては、繰り返し継代しても細胞の形質が変化せずに無限増殖し、他の形質は癌細胞とは区別可能な非ヒト動物細胞であり、培養が簡易かつ高効率でDNAの導入が可能な細胞が好ましい。

癌細胞と区別可能な形質として、上述したヒト癌細胞に特異的な性質と相対する性質、例えば、足場依存性増殖、細胞増殖時の接触阻止、増殖時の血清要求性等をあげることができる。また、後述のように、宿主細胞由来のDNAと、外来のヒト癌細胞由来のDNAを区別して解析する必要があるため、非ヒト動物細胞を宿主細胞に用いることが好ましい。外来のヒトの遺伝子が宿主細胞内で発現し、該遺伝子を含んでいた細胞と同様の機能を有する形質転換細胞を獲得する

ためには、分類上あまりヒトとかけ離れていない動物、好ましくはヒト以外の哺乳動物の細胞をあげることができる。具体的には、マウスの繊維芽細胞株 NIH3T3 (理化学研究所細胞開発銀行 RCB0150)、ラット繊維芽細胞株 Rat-2 [Virology, 113, 408 (1981)、ATCC CRL-1764] をあげることができる。

以上の方法により、形質転換された細胞、すなわち癌細胞由来の原因遺伝子が導入された細胞は、軟寒天培地などの培地中でコロニーを形成することが可能となる。

例えば、軟寒天培地中では生育できない足場依存性増殖を示す繊維芽細胞に、ヒト癌細胞株の染色体DNA由来のDNA断片を導入した後、軟寒天培地中で培養すると、ヒト癌細胞株の有する足場非依存性の形質を獲得した細胞は軟寒天中でコロニーを形成する。該コロニー形成を指標にすることにより、形質転換細胞を容易に単離することが可能となる。

上記で単離されたコロニー形成細胞を用い、常法に従いDNA分析することにより、ヒト癌遺伝子を特定し、取得することができる。すなわち、EcoRI等の適当な制限酵素で該細胞の染色体DNAを切断後、サザンブロット解析 (モレキュラー・クローニング第2版、DNAクローニング1) を行うことよりヒト癌遺伝子を特定し、取得することができる。

サザンブロット解析に用いるプローブとしては、ヒトの遺伝子で特異的な配列として普遍的にヒトゲノム遺伝子内に存在するAlu配列を有する、BLUR8(Proc. Natl. Acad. Sci. USA, 77, 1398 (1980)) などを用いることができるが、任意のヒト細胞から染色体DNAを単離し、放射性同位体、酵素などで標識したDNA断片を用いることもできる。Alu配列は、ヒトゲノムのイントロンに特異的に見られる配列であり、かつヒトゲノムDNA中に30万コピーと他の塩基配列と比較して圧倒的多数を占めるため、標識されたDNAの大部分はAlu配列であると考えられる。従って、標識した染色体DNA断片を、癌遺伝子中のイントロンを特定するためのプローブとして用いることができる。

サザンブロットによる解析の結果、同一の長さのバンドが検出されたコロニーを選択した後、該コロニーより染色体DNAを単離し、該染色体DNA由来のDNA断片を用い、上述の宿主細胞への導入からサザンブロットまでの操作を繰り返す。そして、選択したコロニー由来の染色体DNA断片が、形質転換を引き起こすこと、サザンブロットによる解析で、最初と同じ長さのバンドが検出されることを確認することにより、該DNA断片により、癌形質が誘発されることがわかる。

形質転換された細胞より染色体DNAを単離したのち、EcoRI等の適当な制限酵素で数kb~20kb程度に部分的に切断する。該DNA断片をクローニングベクターに組み込み、宿主細胞に導入することによりDNAライブラリーを作製する。

具体的なDNAライブラリー作製法としては、モレキュラー・クローニング第2版等に記載された方法等があげられる。

DNAライブラリーを作製するためのクローニングベクターとしては、大腸菌K12株中で自律複製できるものであれば、ファージベクター、プラスミドベクター等いずれも使用できる。

具体的には、ZAP Express (Stratagene社製、Strategies, 5, 58 (1992))、pBluescript II SK(+) (Nucleic Acids Research, 17, 9494 (1989))、Lambda ZAP II (Stratagene社製)、 λ gt10、 λ gt11 (DNA Cloning, A Practical Approach, 1, 49 (1985))、 λ TriplEx (Clontech社製)、 λ ExCell (Pharmacia社製)、 λ DASH II (Stratagene社製)、該 λ DASH IIファージベクターを基に構築した λ PSL、pT7T318U (Pharmacia社製)、pcD2 (Mol. Cell. Biol., 3, 280 (1983))等、9~23kbの長いDNAを挿入可能なベクターをあげることができる。

宿主微生物としては、大腸菌Escherichia coliに属する微生物であればいずれも用いることができる。具体的には、Escherichia coli XL1-Blue MRF' (Stratagene社製、Strategies, 5, 81 (1992))、Escherichia coli C600 (Genetics, 39, 440 (1954))、Escherichia coli Y1088 (Science, 222, 778 (1983))、Escherichia coli Y1090 (Science, 222, 778 (1983))、Escherichia coli NM522

[J. Mol. Biol., 166, 1 (1983)]、Escherichia coli K802 [J. Mol. Biol., 16, 118 (1966)]、Escherichia coli JM105 [Gene, 38, 275 (1985)] および Escherichia coli LE392 (モレキュラー・クローニング第2版) 等を用いることができる。

上述のように調製したDNAライブラリーに対し、前述のプロープを用いてコロニー・ハイブリダイゼーションあるいはブランク・ハイブリダイゼーション (モレキュラー・クローニング第2版) を行うことより、該DNAライブラリーより目的とするDNAクローンを取得することができる。

該クローンより得られたヒト癌細胞由来の癌遺伝子を含むDNA断片は、染色体DNA由来であるため、イントロンも含まれていると考えられる。該遺伝子がコードする蛋白質の構造を明らかにするために、該遺伝子のcDNAを、cDNAライブラリーをスクリーニングすることによって取得する。

(2) 新規チロシンフォスファターゼ cDNAの取得

cDNAライブラリーを作製するために、適切な細胞または組織より全RNAあるいはmRNAを調製する。例えば、細胞としてヒト胃癌セルラインMKN45 [Jpn. J. Cancer Res. 77, 849 (1986)、理化学研究所細胞開発銀行 RCB1001] などをあげることができる。

全RNAを調製する方法として、ヒト細胞からチオシアン酸グアニジントリフルオロ酢酸セシウム法 [Methods in Enzymol., 154, 3 (1987)]、酸性チオシアン酸グアニジン・フェノール・クロロホルム法 [Analytical Biochemistry, 162, 156 (1987)] などを用いることができる。

全RNAからポリ(A)⁺RNAとしてmRNAを調製する方法として、オリゴ(dT)セルロース法 (モレキュラー・クローニング第2版) などを用いることができる。

あるいは、Fast Track mRNA Isolation Kit (Invitrogen社)、Quick Prep mRNA Purification Kit (Pharmacia社) 等のキットを用いてヒト細胞から直接mRNAを調製することもできる。

cDNAを調製し、該cDNAを適当なベクターに組み込み、宿主細胞に導入することによりcDNAライブラリーを作製する。具体的なcDNAライブラリー作製法としては、モレキュラー・クローニング第2版等に記載された方法、GublerとHoffmanの方法〔Gene, 25, 263 (1983)〕、あるいは市販のキット、例えばSuperScript Plasmid System for cDNA Synthesis and Plasmid Cloning (ライフ・テクノロジー社製)やZAP-cDNA Synthesis Kit (Staratagene社製)を用いる方法等があげられるが、前述(1)でのDNAライブラリーの作製方法に準じて行うことができる。

調製したcDNAライブラリーに対してコロニー・ハイブリダイゼーションあるいはブラーク・ハイブリダイゼーション(モレキュラー・クローニング第2版)を行うことにより新規ヒト癌遺伝子を含有するcDNAクローンを得ることができる。

プローブとしては、(1)で得られた癌遺伝子を含む染色体DNA断片を用いることができる。この場合ハイブリダイゼーションのバックグラウンドを減らすために、上記の染色体DNA断片から200bp～2 kb程度の長さでエクソンを含む断片を以下のようにして単離してプローブに用いるのが望ましい。

すなわち、上記の染色体DNA断片を適当な制限酵素で切断して200bp～2 kbの長さのDNA断片を単離し、プローブとする。上述(1)で取得したクローンより、上述の方法でRNAまたはmRNAを単離する。単離した該DNA断片および該RNAを用いてノーザン・ブロット・ハイブリダイゼーションを行ない、はっきりとしたバンドが得られた断片をエクソンを含む断片として選択し、プローブとして使用する。

このようにして得られたcDNAクローン中に存在する新規ヒト癌遺伝子のcDNAの塩基配列を、後述(3)の方法により決定することができる。決定された塩基配列より、全長のcDNAが得られていないと考えられる場合には以下の方法で全長cDNAを取得することができる。すなわち、前述の方法で得たcDNAの両端に

アダプターを付加し、このアダプターの塩基配列と得られたcDNAの塩基配列に基づいたプライマーでPCRを行う5'-RACE(rapid amplification of cDNA ends)および3'-RACE〔Proc. Natl. Acad. Sci. USA, 85, 8998 (1988)〕により、プライマーに用いた配列よりも5'末端側および3'末端側のcDNA断片を得ることができる。得られたcDNA断片をcDNAクローンのcDNAと連結させることにより、本発明の全長cDNAを取得することができる。

(3) cDNAの塩基配列の解析

得られたcDNAをそのまま、あるいは該cDNA部分を適当な制限酵素で切断後、pUC118等の適当なクローニングベクターにサブクローニングした後、通常用いられる塩基配列解析方法、例えばSangerらのジデオキシ法〔Proc. Natl. Acad. Sci. USA, 74, 5463 (1977)〕あるいは373A・DNAシーケンサー(Perkin Elmer社製)等の塩基配列分析装置を用いて分析することにより、塩基配列を決定することができる。

決定された塩基配列より、該cDNAがコードする蛋白質のアミノ酸配列を知ることができる。

該塩基配列の新規性に関しては、BLAST等の相同性検索プログラムを用いて、GenBank、EMBLおよびDDBJ等の塩基配列データベースを検索することにより、確認することができる。

上記方法により決定される塩基配列として、HD-PTPをコードするcDNAの有する配列番号1に示される塩基配列をあげることができる。

(4) ゲノムDNAの取得

ゲノムDNAは、ヒトゲノムDNAを鋳型にして、上記(3)で得られたcDNAの塩基配列に基づいて設計・合成できるDNAをプライマーに用いたPCR〔PCR, A practical Approach, Oxford University Press (1991)〕によって取得することができる。また、Clontech社等から購入できるヒトゲノムDNAライブラリーから(3)で得られたcDNAをプローブにしてゲノムDNAクローンを得ることもで

きる。

(5) HD-PTPをコードするDNAの調製

HD-PTPのcDNAの塩基配列に基づいて設計したプライマーDNAを合成し、ヒト組織あるいは細胞から(3)と同様にして調製したcDNAあるいはcDNAライブラリーを鋳型として、PCRを行うことによりHD-PTPをコードするDNA(以下、HD-PTP DNAと略す)を増幅し取得することができる。

また決定されたDNAの塩基配列に基づいて、DNA合成機で化学合成することによってHD-PTP DNAを調製することもできる。DNA合成機としては、フォスフォアミダイト法を利用したPerkin Elmer社製のDNA合成機model 392等をあげることができる。

(6) HD-PTPオリゴヌクレオチドの調製

上述の方法で取得した本発明のDNAおよびDNA断片を用いて、常法あるいは上述のDNA合成機により、本発明のDNAの一部の配列を有するオリゴヌクレオチドあるいは相補的な塩基配列をもつオリゴヌクレオチド(これらを以下、HD-PTPオリゴヌクレオチドと称する)を調製することができる。

HD-PTPオリゴヌクレオチドとしては、具体的には、配列番号1、3、40または41で表される塩基配列中の連続した10~60塩基と同じ配列を有するDNAまたは該DNAと相補的な配列を有するDNAをあげることができる。センスプライマーおよびアンチセンスプライマーとして用いる場合には、両者の融解温度および塩基数が極端に変わることのない上述のオリゴヌクレオチドが好ましい。アンチセンスプライマーとしては、具体的には、配列番号24、25、28または29で表される塩基配列を有するDNAをあげることができる。

さらに、これらオリゴヌクレオチドの誘導体も本発明のオリゴヌクレオチドとしてあげることができる。該オリゴヌクレオチドの誘導体としては、オリゴヌクレオチド中のリン酸ジエステル結合がホスフロチオエート結合に変換されたオリゴヌクレオチド誘導体、オリゴヌクレオチド中のリン酸ジエステル

結合がN 3' - P 5' ホスフォアミデート結合に変換されたオリゴヌクレオチド誘導体、オリゴヌクレオチド中のリボースとリン酸ジエステル結合がペプチド核酸結合に変換されたオリゴヌクレオチド誘導体、オリゴヌクレオチド中のウラシルがC-5 プロビニルウラシルで置換されたオリゴヌクレオチド誘導体、オリゴヌクレオチド中のウラシルがC-5 チアゾールウラシルで置換されたオリゴヌクレオチド誘導体、オリゴヌクレオチド中のシトシンがC-5 プロビニルシトシンで置換されたオリゴヌクレオチド誘導体、オリゴヌクレオチド中のシトシンがフェノキサジン修飾シトシン (phenoxazine-modified cytosine) で置換されたオリゴヌクレオチド誘導体、オリゴヌクレオチド中のリボースが2' - O-プロピルリボースで置換されたオリゴヌクレオチド誘導体、あるいはオリゴヌクレオチド中のリボースが2' - メトキシエトキシリボースで置換されたオリゴヌクレオチド誘導体等をあげることができる〔細胞工学, 16, 1463 (1997)〕。

[2] 遺伝子組換え技術を用いた、HD-PTPの調製

HD-PTPは、モレキュラー・クローニング第2版やカレント・プロトコールズ・イン・モレキュラー・バイオロジー等に記載された方法を用い、例えば以下の方法により、上記[1]に記載の方法により取得したHD-PTP DNAを宿主細胞で発現させて、調製することができる。

HD-PTP DNAを適当な発現ベクターのプロモーター下流に挿入した組換え体DNAを造成する。該組換え体DNAを宿主細胞に導入することにより、HD-PTP蛋白質を発現する形質転換体を得ることができる。

宿主細胞としては、細菌、酵母、動物細胞、昆虫細胞、植物細胞等、目的とする遺伝子を発現できるものであればいずれも用いることができる。

発現ベクターとしては、上記宿主細胞において自律複製可能なしは染色体中への組込が可能で、HD-PTP DNAを転写できる位置にプロモーターを含有しているものが用いられる。

細菌等の原核生物を宿主細胞として用いる場合は、該組換え体DNAは原核生物中で自律複製可能であると同時に、プロモーター、リボソーム結合配列、HD-PTP DNA、転写終結配列、より構成されていることが好ましい。プロモーターを制御する遺伝子が含まれていてもよい。

発現ベクターとしては、例えば、pKK233-2 (Pharmacia社)、pSE280 (Invitrogen社)、pGEMEX-1 (Promega社)、pQE-8 (QIAGEN社)、pKYP10 (特開昭58-110600)、pKYP200 [Agricultural Biological Chemistry, 48, 669 (1984)]、pLSA1 [Agric. Biol. Chem., 53, 277 (1989)]、pGEL1 [Proc. Natl. Acad. Sci. USA, 82, 4306 (1985)]、pBluescript II SK(-) (Stratagene社)、pGEX (Pharmacia社)、pET-3 (Novagen社) 等をあげることができる。

プロモーターとしては、大腸菌や枯草菌等の宿主細胞中で発現できるものであればいかなるものでもよい。例えば、trpプロモーター (P_{trp})、lacプロモーター、 P_L プロモーター、 P_R プロモーター、T7プロモーター等の、大腸菌やファージ等に由来するプロモーターをあげることができる。また P_{trp} を2つ直列させたプロモーター ($P_{trp} \times 2$)、tacプロモーター、lacT7プロモーター、letIプロモーターのように人為的に設計改変されたプロモーター等も用いることができる。

リボソーム結合配列であるShine-Dalgarno配列と開始コドンとの間を適当な距離 (例えば6~18塩基) に調節したプラスミドを用いることが好ましい。

本発明の組換えベクターにおいては、本発明のDNAの発現には転写終結配列は必ずしも必要ではないが、構造遺伝子の直下に転写終結配列を配置することが好ましい。

宿主細胞としては、エシェリヒア属、バチルス属、コリネバクテリウム属、ブレヴィバクテリウム属、シュードモナス属、セラチア属、等に属する微生物、例えば、Escherichia coli XL1-Blue、Escherichia coli XL2-Blue、Escherichia coli DH1、Escherichia coli MC1000、Escherichia coli KY3276、Escherichia coli W1485、Escherichia coli JM109、Escherichia coli HB101、Escherichia

coli No.49、Escherichia coli W3110、Escherichia coli NY49、Serratia ficaria、Serratia fonticola、Serratia liquefaciens、Serratia marcescens、Bacillus subtilis、Bacillus amyloliquefaciens、Brevibacterium ammoniagenes、Brevibacterium immariophilum ATCC14068、Brevibacterium saccharolyticum ATCC14066、Corynebacterium glutamicum ATCC13032、Corynebacterium glutamicum ATCC14067、Corynebacterium glutamicum ATCC13869、Corynebacterium acetoacidophilum ATCC13870、Microbacterium ammoniaphilum ATCC15354、Pseudomonas sp. D-0110等をあげることができる。

組換え体DNAの導入方法としては、前記宿主細胞へDNAを導入する方法であればいずれも用いることができ、例えば、塩化カルシウム法〔Proc. Natl. Acad. Sci. USA, 69, 2110 (1972)〕、プロトプラスト法 (特開昭63-248394)、エレクトロポレーション法〔Gene, 17, 107 (1982)、Molecular & General Genetics, 168, 111 (1979)〕等をあげることができる。

酵母菌株を宿主細胞として用いる場合には、発現ベクターとして、例えば、YEp13 (ATCC37115)、YEp24 (ATCC37051)、YCp50 (ATCC37419)、pHS19、pHS15等を用いることができる。

プロモーターとしては、酵母菌株中で発現できるものであればいずれのものを用いてもよく、例えば、PH05プロモーター、PGKプロモーター、GAPプロモーター、ADHプロモーター、gal 1プロモーター、gal 10プロモーター、ヒートショックポリペプチドプロモーター、MF α 1プロモーター、CUP 1プロモーター等のプロモーターをあげることができる。

宿主細胞としては、サッカロマイセス属、シゾサッカロマイセス属、クルイベロミセス属、トリコスボロン属、シワニオミセス属、ピヒア属等に属する酵母菌株をあげることができる。具体的には、Saccharomyces cerevisiae、Schizosaccharomyces pombe、Kluyveromyces lactis、Trichosporon pullulans、Schwanniomyces alluvius、Pichiapastris等をあげることができる。

組換え体DNAの導入方法としては、酵母にDNAを導入する方法であればいずれも用いることができ、例えば、エレクトロポレーション法〔Methods in Enzymol., 194, 182 (1990)〕、スフェロプラスト法〔Proc. Natl. Acad. Sci. USA, 81, 4889 (1984)〕、酢酸リチウム法〔Journal of Bacteriology, 153, 163 (1983)〕等をあげることができる。

動物細胞を宿主として用いる場合の発現ベクターとしては、宿主動物細胞で転写を行なうプロモーター、HD-PTP DNA、転写の終止と転写物のポリアデニル化のシグナルの配列を含有しているものが用いられる。またベクターの作製や維持を容易にするため、*Escherichia. coli*内でも自律複製と遺伝子導入マーカーとなる薬剤耐性遺伝子を発現できるものが望ましい。

プロモーターとしては、動物細胞中で転写を行なえるものであればいずれも用いることができるが、SV40の初期プロモーター、ヒトサイトメガロウイルス (CMV) のIE (immediate early) 遺伝子のプロモーター、ラウス肉腫ウイルス (Rous sarcoma virus; RSV)、ヒト免疫不全ウイルス (human immunodeficiency virus; HIV)、モロニーマウス白血病ウイルス (Moloney mouse leukemia virus; MMLV) 等レトロウイルスのロング・ターミナル・リピート (long terminal repeat; LTR) などのウイルス由来の配列を有するプロモーター、SV40の初期プロモーターにヒトT細胞白血病ウイルス I のLTRを人為的につなげたSR α プロモーター、またはメタロチオネイン遺伝子や β -アクチン遺伝子、伸長因子 (Elongation factor) - 1 遺伝子などの動物細胞由来の遺伝子のプロモーター等をあげることができる。また、ヒトCMVのIE遺伝子のエンハンサーをプロモーターと共に用いてもよい。

動物細胞を宿主として用いる場合には、発現ベクターとして、例えば、pAGE107〔特開平3-22979、Cytotechnology, 3, 133, (1990)〕、pAS3-3 (特開平2-227075)、pCDM8〔Nature, 329, 840, (1987)〕、pcDNA1/Amp (Invitrogen社製)、pREP4 (Invitrogen社製)、pAGE103〔Journal of Biochemistry, 101, 1307 (1987)〕

〕、pCDL81等が用いられる。pCDL81は、ハイグロマイシン耐性遺伝子およびEBウイルス複製起点をもつベクターEB0-pcD [Mol. Cell. Biol., 8, 2837 (1988)] にSR α プロモーターを組み込み、さらにマルチクローニングサイト (XhoI-NotI-XbaI-KpnI-BamHI) およびPoly Aシグナルの塩基配列を有するDNAを挿入して作製したベクターである。

宿主細胞としては、マウス・ミエローマ細胞、ラット・ミエローマ細胞、マウス・ハイブリドーマ細胞、マウス繊維芽細胞であるNIH3T3細胞、ラット繊維芽細胞であるRat-2細胞、ヒトの細胞であるナマルバ (Namalwa) 細胞、ヒト胎児腎臓細胞、ヒト白血病細胞、アフリカミドリザル腎臓細胞、チャイニーズ・ハムスターの細胞であるCHO細胞、HBT5637 (特開昭63-299) 等をあげることができる。

マウス・ミエローマ細胞としては、SP2/0、NS0等、ラット・ミエローマ細胞としてはYB2/0等、ヒトの細胞であるナマルバ (Namalwa) 細胞としてはNamalwa KJM-1細胞、ヒト胎児腎臓細胞としてはHEK293(ATCC:CRL-1573)、ヒト白血病細胞としてはBALL-1、アフリカミドリザル腎臓細胞としてはCOS-1、COS-7等をあげることができる。

組換え体DNAの導入方法としては、動物細胞にDNAを導入する方法であればいずれも用いることができ、例えば、エレクトロポレーション法 [Cytotechnology, 3, 133 (1990)]、リン酸カルシウム法 (特開平2-227075)、リポフェクション法 [Proc. Natl. Acad. Sci. USA, 84, 7413 (1987)] 等をあげることができる。

宿主染色体DNAにHD-PTP DNAが組み込まれた恒常的なHD-PTP発現細胞は、G418、ハイグロマイシン等の薬剤に対する耐性遺伝子を発現できる配列を含むHD-PTP発現ベクターを宿主細胞に導入し、薬剤の存在下で培養することにより選択することができる。また、宿主細胞中でのHD-PTPの生産量を上昇させるために、ジヒドロ葉酸レダクターゼ (dihydrofolate reductase; dhfr) 遺伝子を発現で

きるような配列を含むHD-PTP恒常的発現ベクターを宿主細胞に導入し、dhfr阻害剤であるメトトレキサート (methotrexate) の濃度を段階的に上げながら培養することにより、dhfr遺伝子とともにHD-PTP DNAのコピー数を増幅させることもできる。このdhfr遺伝子を用いた遺伝子増幅を行なう場合の宿主細胞としては、dhfr遺伝子が機能していない細胞、例えばCHO/dhfr⁻ (ATCC:CRL-9096) などを用いる。

昆虫細胞を宿主として用いる場合には、例えばBaculovirus Expression Vectors, A Laboratory Manual, W. H. Freeman and Company, New York (1992) (以下、バキュロウイルス・イクスプレッション・ベクターズ・ア・ラボラトリー・マニュアルと略す)、Bio/Technology, 6, 47 (1988)、カレント・プロトコールズ・イン・モレキュラー・バイオロジー等に記載された方法によって、ポリペプチドを発現することができる。

すなわち、組換え遺伝子導入ベクターおよびバキュロウイルスを昆虫細胞に共導入して昆虫細胞培養上清中に組換えウイルスを得た後、さらに組換えウイルスを昆虫細胞に感染させることにより、蛋白質を発現させることができる。

該方法において用いられる遺伝子導入ベクターとしては、例えば、pVL1392、pVL1393、pBlueBacIII (ともにInvitrogen社製) 等をあげることができる。

バキュロウイルスとしては、例えば、夜盗蛾科昆虫に感染するウイルスであるアウトグラフィア・カリフォルニカ・ヌクレアー・ポリヘドロシス・ウイルス (Autographa californica nuclear polyhedrosis virus) などを用いることができる。

昆虫細胞としては、Spodoptera frugiperdaの卵巣細胞であるSf9、Sf21 (バキュロウイルス・イクスプレッション・ベクターズ・ア・ラボラトリー・マニュアル)、Trichoplusia niの卵巣細胞であるHigh 5 (Invitrogen社製) 等を用いることができる。

組換えウイルスを調製するための、昆虫細胞への上記組換え遺伝子導入ベク

ターと上記バキュロウイルスの共導入方法としては、例えば、リン酸カルシウム法（特開平2-227075）、リポフェクション法〔Proc. Natl. Acad. Sci. USA, 84, 7413 (1987)〕等をあげることができる。

遺伝子の発現方法としては、直接発現以外に、モレキュラー・クローニング第2版に記載されている方法等に準じて、分泌生産、融合蛋白質発現等を行うことができる。

酵母、動物細胞、昆虫細胞、植物細胞または植物体で発現させた場合には、糖あるいは糖鎖が付加された蛋白質を得ることができる。

以上のようにして得られる形質転換体を培地に培養し、培養物中に本発明の蛋白質を生成蓄積させ、該培養物から採取することにより、本発明の蛋白質を製造することができる。

また、患者の生体内から採取した細胞に、適切な本発明の蛋白質を発現させるための発現ベクターを導入した後、細胞を生体内に戻すことにより、本発明の蛋白質を患者の生体内に発現させることもできる。

本発明の形質転換体を培地に培養する方法は、宿主の培養に用いられる通常の方法に従って行うことができる。

大腸菌等の原核生物あるいは酵母等の真核生物を宿主として得られた形質転換体を培養する培地としては、該生物が資化し得る炭素源、窒素源、無機塩類等を含有し、形質転換体の培養を効率的に行える培地であれば天然培地、合成培地のいずれを用いてもよい。

炭素源としては、該生物が資化し得るものであればよく、グルコース、フラクトース、スクロース、これらを含有する糖蜜、デンプンあるいはデンプン加水分解物等の炭水化物、酢酸、プロピオン酸等の有機酸、エタノール、プロパノールなどのアルコール類等を用いることができる。

窒素源としては、アンモニア、塩化アンモニウム、硫酸アンモニウム、酢酸アンモニウム、リン酸アンモニウム等の無機酸もしくは有機酸のアンモニウム

塩、その他の含窒素化合物、並びに、ペプトン、肉エキス、酵母エキス、コーンステープリカー、カゼイン加水分解物、大豆粕および大豆粕加水分解物、各種発酵菌体、およびその消化物等を用いることができる。

無機物としては、リン酸第一カリウム、リン酸第二カリウム、リン酸マグネシウム、硫酸マグネシウム、塩化ナトリウム、硫酸第一鉄、硫酸マンガン、硫酸銅、炭酸カルシウム等を用いることができる。

培養は、通常振盪培養または深部通気攪拌培養などの好氣的条件下で行う。培養温度は15～40℃がよく、培養時間は、通常16～96時間である。培養中pHは3.0～9.0に保持する。pHの調整は、無機または有機の酸、アルカリ溶液、尿素、炭酸カルシウム、アンモニアなどを用いて行う。

また、培養中に必要に応じて、アンピシリンやテトラサイクリン等の抗生物質を培地に添加してもよい。

プロモーターとして誘導性のプロモーターを用いた発現ベクターで形質転換した微生物を培養するときには、必要に応じてインデューサーを培地に添加してもよい。例えば、lacプロモーターを用いた発現ベクターで形質転換した微生物を培養するときにはイソプロピルーβ-D-チオガラクトピラノシド等を、trpプロモーターを用いた発現ベクターで形質転換した微生物を培養するときにはインドールアクリル酸等を培地に添加してもよい。

動物細胞を宿主として得られた形質転換体を培養する培地としては、一般に使用されているRPMI1640培地〔The Journal of the American Medical Association, 199, 519 (1967)〕、EagleのMEM培地〔Science, 122, 501 (1952)〕、DMEM培地〔Virology, 8, 396 (1959)〕、199培地〔Proceeding of the Society for the Biological Medicine, 73, 1 (1950)〕またはこれら培地に牛胎児血清等を添加した培地等を用いることができる。

培養は、通常5%CO₂存在下等の条件下で行う。培養温度は35～37℃がよく、培養時間は、通常3～7日間である。

また、培養中に必要に応じて、カナマイシン、ペニシリン等の抗生物質を培地に添加してもよい。

昆虫細胞を宿主として得られた形質転換体を培養する培地としては、一般に使用されているTNM-FH培地（Pharmingen社製）、Sf-900 II SFM培地（LifeTechnologies社製）、ExCell400、ExCell405（いずれもJRH Biosciences社製）等を用いることができる。

培養温度は25～30℃がよく、培養時間は、通常1～4日間である。

また、培養中に必要に応じて、ゲンタマイシン等の抗生物質を培地に添加してもよい。

上記形質転換体の培養液から、上記方法により発現させた蛋白質を単離精製するためには、通常の蛋白質の単離、精製法を用いればよい。

例えば、本発明の蛋白質が、細胞内に溶解状態で発現した場合には、培養終了後、細胞を遠心分離により回収し水系緩衝液にけん濁後、超音波破碎機、フレンチプレス、マントンガウリンホモゲナイザー、ダイノミル等により細胞を破碎し、無細胞抽出液を得る。該無細胞抽出液を遠心分離することにより得られた上清から、通常の蛋白質の単離精製法、即ち、溶媒抽出法、硫酸等による塩析法、脱塩法、有機溶媒による沈殿法、ジエチルアミノエチル（DEAE）セファロース、DIAION HPA-75（三菱化学社製）等のレジンを用いた陰イオン交換クロマトグラフィー法、S-Sepharose FF（Pharmacia社製）等のレジンを用いた陽イオン交換クロマトグラフィー法、ブチルセファロース、フェニルセファロース等のレジンを用いた疎水性クロマトグラフィー法、分子篩を用いたゲルろ過法、アフィニティークロマトグラフィー法、クロマトフォーカシング法、等電点電気泳動等の電気泳動法等の手法を単独あるいは組み合わせて用い、精製標品を得ることができる。

また、該蛋白質が細胞内に不溶体を形成して発現した場合は、同様に細胞を回収後破碎し、遠心分離を行うことにより得られた沈殿画分として該蛋白質の

不溶体を回収後、該蛋白質の不溶体を蛋白質変性剤で可溶化する。該可溶化液を、蛋白質変性剤を含まないあるいは蛋白質変性剤の濃度が蛋白質が変性しない程度に希薄な溶液に希釈、あるいは透析し、該蛋白質を正常な立体構造に還元させた後、上記と同様の単離精製法により精製標品を得ることができる。

本発明の蛋白質あるいはその糖修飾体等の誘導体が細胞外に分泌された場合には、培養上清に該蛋白質あるいはその糖鎖付加体等の誘導体を回収することができる。即ち、該培養物を上記と同様の遠心分離等の手法により処理することにより培養上清を取得し、該培養上清から、上記と同様の単離精製法を用いることにより、精製標品を得ることができる。

また、本発明の蛋白質は、Fmoc法（フルオレニルメチルオキシカルボニル法）、tBoc法（t-ブチルオキシカルボニル法）等の化学合成法によっても製造することができる。また、Advanced ChemTech社、Perkin Elmer社、Pharmacia社、Protein Technology Instrument社、Synthecell-Vega社、PerSeptive社、島津製作所等のペプチド合成機を利用し化学合成することもできる。

精製した本発明の蛋白質の構造解析は、蛋白質化学で通常用いられる方法、例えば遺伝子クローニングのためのタンパク質構造解析（平野久著、東京化学同人発行、1993年）に記載の方法により実施可能である。

[3] HD-PTP蛋白質を認識する抗体の調製

・抗体産生株の作製

上記[2]に記載の方法により取得したHD-PTP蛋白質の全長または部分断片の精製標品、あるいはHD-PTP蛋白質の部分ペプチド（ペプチド合成機を利用し化学合成できる）を抗原として免疫する。免疫する方法としては、動物の皮下、静脈内または腹腔内に抗原をそのまま投与してもよいが、抗原性の高いキャリアタンパク質を結合させて投与したり、あるいは適当なアジュバントとともに抗原を投与することが好ましい。

キャリアタンパク質としては、スカシガイヘモシアニン、キーホールリンペ

ットヘモシアニン、牛血清アルブミン、牛チログロブリン等があげられ、アジュバンドとしては、フロインドの完全アジュバント(Complete Freund's Adjuvant)、水酸化アルミニウムゲルと百日咳菌ワクチン等があげられる。

免疫動物としては、ウサギ、ヤギ、3～20週令のマウス、ラット、ハムスターなどの非ヒト哺乳動物があげられる。

抗原の投与は、1回目の投与の後、1～2週間毎に3～10回行う。抗原の投与量は動物1匹当たり50～100 μ gが好ましい。各投与後、3～7日目に免疫動物の眼底静脈叢あるいは尾静脈より採血し、該血清の抗原との反応性について、酵素免疫測定法〔酵素免疫測定法(ELISA法)：医学書院刊(1976年)〕などで確認する。

そして、該血清が十分な抗体価を示した非ヒト哺乳動物を、血清または抗体産生細胞の供給源とする。

ポリクローナル抗体は、該血清を分離、精製することにより調製することができる。

モノクローナル抗体は、該抗体産生細胞と非ヒト哺乳動物由来の骨髓腫細胞とを融合させてハイブリドーマを作製し、該ハイブリドーマを培養するか、動物に投与して該細胞を腹水癌化させ、該培養液または腹水を分離、精製することにより調製することができる。

抗体産生細胞は、抗原投与された非ヒト哺乳動物脾細胞、リンパ節、末梢血などから採取する。

骨髓腫細胞としては、マウスから得られた株化細胞である、8-アザグアニン耐性マウス(BALB/c由来)骨髓腫細胞株P3-X63Ag8-U1(P3-U1)〔Europ. J. Immunol., 6, 511 (1976)〕、SP2/0-Ag14(SP-2)〔Nature, 276, 269 (1978)〕、P3-X63-Ag8653(653)〔J. Immunol., 123, 1548 (1979)〕、P3-X63-Ag8(X63)〔Nature, 256, 495 (1975)〕など、イン・ビトロ (*in vitro*) で増殖可能な骨髓腫細胞であればいかなるものでもよい。これらの細胞株の培養および継代につ

いてはアンチボディース・ア・ラボラトリー・マニュアル [Antibodies -A Laboratory Manual, Cold Spring Harbor Laboratory, (1988)、以下、アンチボディース・ア・ラボラトリー・マニュアルと略す] に従い、細胞融合時までに 2×10^7 個以上の細胞数を確保する。

上記で得られた抗体産生細胞と骨髓腫細胞とを洗浄したのち、ポリエチレングライコール-1000(PEG-1000)などの細胞凝集性媒体を加え、細胞を融合させ、培地中に懸濁させる。細胞の洗浄にはMEM培地またはPBS (リン酸二ナトリウム1.83g、リン酸一カリウム0.21g、食塩7.65g、蒸留水1 リットル、pH7.2) などを用いる。また、融合細胞を懸濁させる培地としては、目的の融合細胞のみを選択的に得られるように、HAT培地 [正常培地 [RPMI-1640 培地に1.5 mmol/Lグルタミン、 5×10^{-5} mol/L 2-メルカプトエタノール、10 μ g/mlジェンタマイシンおよび、10% 牛胎児血清(FCS) (CSL社製) を加えた培地] に 10^{-4} mol/L ヒポキサンチン、 1.5×10^{-5} mol/L チミジンおよび 4×10^{-7} mol/L アミノプテリンを加えた培地] を用いる。

培養後、培養上清の一部をとり、酵素免疫測定法により、抗原蛋白質に反応し、非抗原蛋白質に反応しないサンプルを選択する。ついで、限界希釈法によりクローニングを行い、酵素免疫測定法により安定して高い抗体価の認められたものをモノクローナル抗体産生ハイブリドーマ株として選択する。

・酵素免疫測定法

抗原蛋白質あるいは抗原蛋白質を発現した細胞などを96ウェルプレートにコートし、ハイブリドーマ培養上清もしくは上述の方法で得られる精製抗体を第一抗体として反応させる。

第一抗体反応後、プレートを洗浄して第二抗体を添加する。

第二抗体とは、第一抗体のイムノグロブリンを認識できる抗体を、ビオチン、酵素、化学発光物質あるいは放射線化合物等で標識した抗体である。具体的にはハイブリドーマ作製の際にマウスを用いたのであれば、第二抗体としては、

マウスイムノグロブリンを認識できる抗体を用いる。

反応後、第二抗体を標識した物質に応じた反応を行ない、抗原に特異的に反応するモノクローナル抗体を生産するハイブリドーマとして選択する。

モノクローナル抗体は、ハイブリドーマ細胞を培養して得られる培養液、またはプリスタン処理〔2,6,10,14-テトラメチルペンタデカン(Pristane)0.5mlを腹腔内投与し、2週間飼育する〕した8～10週令のマウスまたはヌードマウスに、モノクローナル抗体産生ハイブリドーマ細胞を腹腔内投与して腹水癌化させた腹水から、分離、精製することにより調製できる。

モノクローナル抗体を分離、精製する方法としては、遠心分離、40～50% 飽和硫酸アンモニウムによる塩析、カプリル酸沈殿法、DEAE-セファロースカラム、陰イオン交換カラム、プロテインAまたはG- カラムあるいはゲル濾過カラム等を用いるクロマトグラフィー等を、単独または組み合わせて行う方法があげられる。この方法により、IgG あるいはIgM 画分を回収し、精製モノクローナル抗体を取得することができる。

精製モノクローナル抗体のサブクラスの決定は、モノクローナル抗体タイピングキットなどを用いて行うことができる。蛋白質量は、ローリー法あるいは280nmでの吸光度より算出することができる。

なお、抗体のサブクラスとは、クラス内のアイソタイプのことで、マウスでは、IgG1、IgG2a、IgG2b、IgG3、ヒトでは、IgG1、IgG2、IgG3、IgG4があげられる。

[4] HD-PTP、HD-PTP DNA、HD-PTPを認識する抗体の利用

(1) [1] に記載したHD-PTP DNAをプローブに用いたノーザンブロット・ハイブリダイゼーション法、HD-PTP DNAの一部の塩基配列を基にしたオリゴヌクレオチドをプライマーに用いたRT-PCR法等を行うことによりHD-PTPのmRNAを検出または定量することができる。該mRNAを検出または定量することにより、HD-PTP遺伝子の発現している組織や細胞を検出すること、HD-PTP遺伝子発現の誘導ま

たは抑制等の情報の取得ができる。したがって、これらのDNAはHD-PTPの発現を測定し、癌の診断薬、HD-PTPの研究用試薬として用いることができる。

(2) [1]記載のHD-PTP DNAあるいはHD-PTPオリゴヌクレオチドを用いてHD-PTP遺伝子の欠失、コピー数の変化、染色体転座等の異常および該遺伝子の塩基配列の置換、欠失、付加等の変異を検出することができる。

HD-PTP遺伝子の欠失、コピー数の変化、染色体転座等の異常の検出方法としては、サザン・ハイブリダイゼーション法があげられる。すなわち、HD-PTP DNAをプローブにして、適当な制限酵素で切断した染色体DNAをサザン・ハイブリダイゼーションすることにより、HD-PTP遺伝子の欠失、コピー数の変化、染色体転座等の異常を確認することができる。

HD-PTP遺伝子の塩基配列の置換、欠失、付加等の変異の検出方法としては、サザン・ハイブリダイゼーション法、PCR法、SSCP (single-strand conformation polymorphism) 法 [Proc. Natl. Acad. Sci. USA, 86, 2766 (1989)] などを用いた方法があげられる。

癌細胞で共通して見出される変異が存在する場合は、この変異部位とハイブリダイズするオリゴヌクレオチドプローブにより染色体DNAをサザン・ハイブリダイゼーション解析することにより、癌の診断を行なうことができる。

(3) [1]記載のHD-PTPのDNAを用いて、ラジエーション・ハイブリッド法 [Science, 250, 245 (1990)] や in situ ハイブリダイゼーション法 [Annals of Human Genetics, 45, 135 (1981)、Cell, 52, 51 (1988)] により、HD-PTP遺伝子の染色体上の位置を決定することができる。

ラジエーション・ハイブリッド法とは、Gene-Bridge 4などのヒト染色体断片をもつ多数のパネル（どの部分の染色体断片が含まれているか染色体マーカーにより解析されているもの）DNAに対し、HD-PTP遺伝子を特異的に増幅させるポリメラーゼ・チェーン・リアクション (Polymerase Chain Reaction、以下、PCRと略す) を行い、増幅結果を解析することにより詳細な染色体の位置を特定す

る方法である。

in situハイブリダイゼーション法では、まず、ヒト染色体の標本に対し、HD-PTPのDNAをプローブとしてハイブリダイズしたシグナルを検出し、標本上のシグナルの位置を特定する。これにより、HD-PTP遺伝子の存在する染色体の番号だけでなく、その染色体上での物理的な位置を特定することができる。プローブは、放射性同位体³Hやビオチンにより標識し、³H標識ではオートラジオグラフにより、ビオチン標識では蛍光色素FITC（フルオロセインイソチオシアネート）で標識したアビジンを用いてシグナルを検出することができる。

(4) [1]記載のHD-PTP DNAを用いて、[2]に記載した方法により、HD-PTPを生産し取得することができる。

(5) HD-PTP遺伝子の異常は発癌に関与しており、該異常遺伝子のコードする蛋白質が発癌に起因していると考えられているため、正常なHD-PTPを投与することにより、癌の治療薬として利用することができる。

本発明のHD-PTPを含有する医薬は、治療薬として単独で投与することも可能ではあるが、通常は薬理学的に許容される一つあるいはそれ以上の担体と一緒に混合し、製剤学の技術分野においてよく知られる任意の方法により製造した医薬製剤として提供するのが望ましい。

投与経路は、治療に際して最も効果的なものを使用するのが望ましく、経口投与、または口腔内、気道内、直腸内、皮下、筋肉内および静脈内等の非経口投与をあげることができ、抗体製剤の場合、望ましくは静脈内投与をあげることができる。

投与形態としては、噴霧剤、カプセル剤、錠剤、顆粒剤、シロップ剤、乳剤、座剤、注射剤、軟膏、テープ剤等があげられる。

経口投与に適当な製剤としては、乳剤、シロップ剤、カプセル剤、錠剤、散剤、顆粒剤等があげられる。

乳剤およびシロップ剤のような液体調製物は、水、ショ糖、ソルビトール、

果糖等の糖類、ポリエチレングリコール、プロピレングリコール等のグリコール類、ごま油、オリーブ油、大豆油等の油類、p-ヒドロキシ安息香酸エステル類等の防腐剤、ストロベリーフレーバー、ペパーミント等のフレーバー類等を添加剤として用いて製造できる。

カプセル剤、錠剤、散剤、顆粒剤等は、乳糖、ブドウ糖、ショ糖、マンニトール等の賦形剤、デンプン、アルギン酸ナトリウム等の崩壊剤、ステアリン酸マグネシウム、タルク等の滑沢剤、ポリビニルアルコール、ヒドロキシプロピルセルロース、ゼラチン等の結合剤、脂肪酸エステル等の界面活性剤、グリセリン等の可塑剤等を添加剤として用いて製造できる。

非経口投与に適切な製剤としては、注射剤、座剤、噴霧剤等があげられる。

注射剤は、塩溶液、ブドウ糖溶液、あるいは両者の混合物からなる担体等を用いて調製される。

座剤はカカオ脂、水素化脂肪またはカルボン酸等の担体を用いて調製される。

また、噴霧剤は該化合物そのもの、ないしは受容者の口腔および気道粘膜を刺激せず、かつ該化合物を微細な粒子として分散させ吸収を容易にさせる担体等を用いて調製される。

担体として具体的には乳糖、グリセリン等が例示される。該化合物および用いる担体の性質により、エアロゾル、ドライパウダー等の製剤が可能である。また、これらの非経口剤においても経口剤で添加剤として例示した成分を添加することもできる。

投与量または投与回数は、目的とする治療効果、投与方法、治療期間、年齢、体重等により異なるが、通常成人1日当たり $10\mu\text{g}/\text{kg} \sim 8\text{mg}/\text{kg}$ である。

(6) (5)においてHD-PTPを外から投与するかわりに、[1]記載のHD-PTP DNAを組み込んだ遺伝子治療用のベクターを患者に投与し、ターゲットとなる細胞内でHD-PTP DNAを発現させることにより、癌の治療を行うことができる。

- (7) [2]記載のHD-PTPを抗原として用い、[3]記載の方法によりHD-PTPを認識する抗体を製造することができる。
- (8) [3]記載の抗体を用いてHD-PTPを検出することができる。具体的にはマイクロタイタープレートを用いるELISA法・蛍光抗体法、ウェスタンブロット法、免疫組織染色等を用いた検出法をあげることができる。
- (9) [3]記載の抗体を用いてHD-PTPを定量することができる。具体的には、液中でHD-PTPと反応する抗体のうち、エピトープが異なる2種類のモノクローナル抗体を用いたサンドイッチELISA法、¹²⁵I等の放射性同位体で標識したHD-PTPとHD-PTPを認識する抗体を用いるラジオイムノアッセイ法等をあげることができる。

図面の簡単な説明

第1図 細胞株KAL1の染色体DNAクローン入SA2-26の挿入DNA断片の制限酵素地図を示す図である。A : AvaI、B : BamHI、E : EcoRI、Rv : EcoRV、S : SalI、X : XhoIの各制限酵素サイトの位置を示す。図の内部の数字は、各EcoRI断片の長さをkbで表わしたものであり、灰色の断片はAluIプローブとハイブリダイズした断片を示す。図の下部の線は、KAL1のmRNAのノーザン・ブロット・ハイブリダイゼーションのプローブに用いた、10種類の断片の位置を示す。*をつけた断片は、KAL1のmRNAとハイブリダイズし約6kbのバンドを検出した、エクソンを含むと考えられる断片である。

第2図 MKN45のcDNAライブラリーの作製に用いた入ファージベクター入PSL1の構造およびcDNAの挿入位置を示す図である。H : HindIII、N : NotIの制限酵素サイト、A-Jは入ファージの遺伝子AからJに位置する遺伝子群を示す。KH54とnin5は入ファージベクターに用いられる遺伝子マーカーである。下に入PSL1の作製に用いたプラスミドベクターpCDL81の構造を示した。SR α : SR α プロモーター、PolyA : polyAシグナル、hph : ハイグロマイシン耐性遺伝子、Ap : アンピシリン

耐性遺伝子を示す。

第3図 HD-PTPと他の蛋白質チロシンフォスファターゼ(PTP) のチロシンフォスファターゼドメインのアミノ酸配列の比較を示す図である。

HD-PTPのチロシンフォスファターゼドメイン(配列番号2の1212~1331番めに相当)のアミノ酸配列と同じアミノ酸を*で示した。+は、各PTPで保存されている活性中心付近のアミノ酸配列の位置を示す。比較した13種類のPTP中12以上のPTPで保存されているにもかかわらず、HD-PTPでは保存されていないアミノ酸を#で示した。比較したPTPの蛋白質データベースSWISS-PROTのアクセス番号(ただしPTP-Hのみは蛋白質データベースPIRのアクセス番号である)を以下に示す。カッコ内は、各アミノ酸配列中のチロシンフォスファターゼドメインの位置である。PTP-1B:P18031(35~279)、PTP-H1:P26045(665~903)、MEG1:P29074(674~913)、PTP- α :P18433(260~503)、PTP- β :P23467(1722~1965)、PTP- δ :P23468(1375~1614)、PTP- ϵ :P23469(154~396)、LAR:P10586(1360~1599)、PTP- γ :P23470(869~1121)、PTP- ζ :P23471(1744~1993)、PTP-2C:Q06124(268~523)、PTP-H:A49724(841~1083)、CD45:P08575(670~912)。

第4図 HD-PTPとラットPTP-TD14のアミノ酸配列の比較を示す図である。

上段がHD-PTPのアミノ酸配列であり、下段がラットPTP-TD14のアミノ酸配列で一致しているアミノ酸は*で示し、一致していないアミノ酸は表記した。-は対応するアミノ酸配列がないことを示す。数字は各蛋白質のN末端からの位置を示す。

第5図 HD-PTP蛋白質の構造を示す模式図である。

AはHD-PTP蛋白質の一次構造上の特徴を示した図で、NがN末端側、CがC末端側である。下の数字はアミノ酸配列の番号を示す。Yはチロシンキナーゼによりリン酸化を受ける可能性のあるチロシン残基の位置、SはセリンスレオニンキナーゼであるMAPキナーゼによりリン酸化を受ける可能性のあるアミノ酸残基の位置、SHBはSH3と結合する可能性のあるモチーフを示す。ヒスチジンド

メイン内の線は、Zn-leaf様構造を取るための繰り返し構造を示す。

BはHD-PTP蛋白質のZn-leaf様構造を含む模式図であり、HおよびCはヒスチジンドメイン中で金属イオンが配位しZn-leaf様構造をとるための、ヒスチジンおよびシステイン残基を表わす。Yはチロシンキナーゼによりリン酸化を受ける可能性のあるチロシン残基、SHBはSH3と結合する可能性のあるモチーフを示す。

第6図 HD-PTPとBR01のアミノ酸配列の比較を示す図である。上段がHD-PTPのアミノ酸配列で下段がBR01のアミノ酸配列を示し、一致しているアミノ酸は|で、類似のアミノ酸は・で示した。数字は各蛋白質のN末端からの位置を示す。

発明を実施するための最良の形態

以下、実施例を示す。

実施例1 HD-PTPのcDNAのクローン化

(1) プロンプとなるDNA断片のクローン化

ヒトB細胞リンパ腫セルラインKAL-1 [Cancer Res., 51, 5392 (1991)] から、プロテイナーゼK/フェノールクロロホルム抽出法 (モレキュラー・クローニング第2版) によりヒト染色体DNAを単離し、該DNAをエタノール沈殿 (モレキュラー・クローニング第2版) により精製した。

ラット繊維芽細胞系セルラインRat-2 [Virology, 113, 408 (1981)、ATCC CRL-1764] を5%の仔ウシ血清を含むダルベッコ改変MEM培地で、10cmディッシュ3枚に培養した。

培養後、ディッシュ1枚あたり、取得されたKAL-1のヒト染色体DNA 30 μ gおよびDNA導入マーカーとなるハイグロマイシン耐性遺伝子をもつプラスミドpHyg [Mol. Cel. Biol., 5, 410 (1985)] のDNA 500ngとを、リン酸カルシウム法により10cmディッシュ内の細胞に共導入させた。

導入後、32時間培養し、得られたディッシュ1枚当りの培養細胞を8枚の培養ディッシュにまきかえた (計24枚)。該細胞を16時間培養した後、培地に終

濃度250 μ g/mlのハイグロマイシンを添加して培養を続けた。DNAが導入された細胞はハイグロマイシン耐性となり、ハイグロマイシン含有培地で増殖し、コロニーを形成する。該培養の結果、約400個のコロニーが形成された。

上記24枚の培養ディッシュで、コロニーを形成した細胞を各ディッシュごとにまとめて回収した。これら回収した細胞を0.5%軟寒天培地（0.5%の寒天および10%のウシ胎児血清を含むDMEM培地）の上にさらに寒天濃度0.33%の軟寒天培地を上層した培地の上に載せて3週間培養した。コロニーを形成した細胞について顕微鏡観察を行い、50細胞以上の細胞数のコロニーを形成した細胞を単離した。

単離した細胞について、上述の方法を用いて染色体DNAを抽出した。該DNA 10 μ gをEcoRIで切断後、アガロースゲル電気泳動を行った。泳動後、ゲル中のDNAをフィルターに転写し、KAL-1の染色体DNAをランダムプライマー法により³²Pで標識したものをヒトAlu配列に対応するプローブ（以下、Alu配列プローブという）を用いてサザン・ハイブリダイゼーションを行い、コロニーに含まれるヒトDNAの分析を行った。

その結果、独立した4個のコロニー由来のDNAが、該プローブとハイブリダイズする7kbのDNA断片を有することがわかった。得られた4つのコロニーから2つを選び、該コロニー由来の染色体DNAを単離・精製した。

再度、上記と同様の操作を繰り返した。すなわち、該DNAをプラスミドpHygとRat-2細胞に共導入し、ハイグロマイシン耐性コロニーを選択した。取得されたコロニーから同様にして染色体DNAを抽出し、抽出したDNA 10 μ gをEcoRIで切断後、アガロースゲル電気泳動を行い、ゲル中のDNAをフィルターに転写後、ヒトAlu配列プローブを用いてサザン・ハイブリダイゼーションを行い、最初に得られた4個のコロニー由来のDNAと同様にハイブリダイズする、7kbのEcoRI断片を有する6個の独立したコロニーを選択した。

該コロニー由来の染色体DNAをEcoRIで部分的に切断し、該DNA断片と、入ファ

ージベクターであるλDASH II (Stratagene社) のEcoRI切断アームとライゲーションし、インビトロパッケージングを行った後、大腸菌株LE392〔モレキュラー・クローニング第2版〕に導入し、DNAライブラリーを作製した。

6×10⁵ブランクについて、上記と同じヒトAlu配列プローブを用いてブランク・ハイブリダイゼーションによりスクリーニングし、約20kbの挿入DNAを持つポジティブクローン2個を単離した。

これら2つのポジティブクロンの挿入DNAを、EcoRIで切断後、アガロースゲル電気泳動を行い、ゲル中のDNAをフィルターに転写後、ヒトAlu配列プローブを用いてサザン・ハイブリダイゼーションを行った。

2つのポジティブクローンとも、挿入DNAは、EcoRIで7.2kb、7.0kb、3.1kb、1.5kb、1.0kb、0.2kbの6断片に切断され、3断片(7.2kb、7.0kb、0.2kb)がAlu配列プローブとハイブリダイズし、残りの3断片(3.1kb、1.0kb、1.5kb)はAlu配列プローブとハイブリダイズしなかった。

さらに、これらポジティブクロンの1つ(λSA2-26)について、第1図に示す制限酵素地図を作成した。各制限酵素で切断したDNA断片について、Alu配列プローブでサザン・プロット・ハイブリダイゼーションを行った。その結果、下記10種類の断片はAlu配列プローブとハイブリダイズせず、Alu配列を含んでいないと判断された。

#6 : 1.0kb EcoRI-EcoRI断片

#7 : EcoRI-BamHI断片

#8 : BamHI-XbaI断片

#9 : XbaI-EcoRV断片

#10 : XbaI-SmaI断片

#11 : XbaI-EcoRI断片

#12 : 1.5kb EcoRI-EcoRI断片

#13 : EcoRI-ApaI断片

#15 : ApaI-BamHI断片

#16 : BamHI-EcoRI断片

上記10種類のDNA断片をそれぞれプローブにして、KAL-1のmRNAに対してノーザン・ブロット・ハイブリダイゼーションを行い、KAL-1のmRNAとハイブリダイズするDNA断片をスクリーニングした。その結果6種類の断片（#6、#8、#11、#13、#15、#16）が6kbのmRNAとハイブリダイズし、エクソンを含む断片であると推定された。これらのDNA断片のうち、#13（0.5kb）と#15（1.0kb）を選択し、下記に述べるcDNAクローニングのプローブとして用いた。

(2) HD-PTP cDNAのクローン化

ヒト胃癌セルラインMKN45 [Jpn. J. Cancer Res. 77, 849 (1986)、理化学研究所細胞開発銀行 RCB1001] のcDNAライブラリーをGublerとHoffmanの方法 [Gene, 25, 263 (1983)] に基づいて作製した。

ベクターとして、 λ PSL1を用いた。 λ PSL1は以下の方法で構築した。

実施例7に後述したpCDL81を制限酵素HindIIIで切断した。得られた約7.9kbのDNA断片2コピーを連結し、該DNAを λ DASH II (Stratagene社製) のHindIII サイト間に挿入し、ベクター λ PSL1を構築した。

λ PSL1の構造を第2図に示した。さらに、第2図にあるように、 λ PSL1をNotIで切断し、pCDL81を1コピー相当分除いたNotI切断DNA断片にcDNAを挿入し、該プラスミドを宿主大腸菌株LE392に挿入してcDNAライブラリーを作製した。この方法は、ベクターに0.5~13kbのDNAが挿入されない場合には、ほとんどファージの生育がみられないため、短いcDNAのクローンを除くことができる。

(1)で得られたDNA断片#13および#15をプローブにして、上述のcDNAライブラリーについて常法（モレキュラー・クローニング第2版）に従ってブランク・ハイブリダイゼーションを行った。その結果、2つのポジティブクローン [cKAL11 (4 kb) およびcKAL16 (5.4kb)] が得られた。

これら2つのクローンのcDNAの制限酵素地図を作成し、構造が異なることが

推定されたため、両者の塩基配列をdye terminator cycle sequencing法により決定し、比較した。

cKAL16の塩基配列を配列番号4に示した。該塩基配列の塩基番号1～3759番目に、1253アミノ酸をコードするオープン・リーディング・フレーム (ORF) が存在していた。

cKAL11の塩基配列を配列番号5に示した。cKAL11は、cKAL16の塩基配列にイントロンが挿入された構造を有していた。すなわち、配列番号5の塩基番号141～529番目、676～960番目、1273～1347番目、1655～1740番目、3680～3773番目、3959～4084番目、4190～4273番目、4413～4504番目、4619～4696番目にイントロンが挿入されていた。

したがってcKAL11はcKAL16と同じ遺伝子由来だが、イントロンが残ったままの不完全なスプライシングのmRNAから合成されたcDNAクローンであると考えられる。

該cKAL16のcDNAの塩基配列の新規性を、塩基配列データベースGenBankに対し解析プログラムBLAST [J. Mol. Biol., 215, 403 (1990)] を用いて検索した結果、新規な塩基配列であることが判明した。また、ラットで最近報告された蛋白質チロシンフォスファターゼ (protein tyrosine phosphatase; 以下、PTPと略す) であるPTP-TD14 [J. Biol. Chem., 273, 21077 (1998)] のcDNAと全体にわたって高い相同性を示し、新規なヒトのPTPをコードするcDNAであることが推測された。cKAL16のcDNA中のORFのアミノ酸配列のC末端側829～1071番目の領域には、図3に示すように他のPTPのフォスファターゼ領域と相同性をもつ領域が存在し、新規PTPをコードしているcDNAと考えられた。しかし、cKAL16のORFはcDNAの5'末端から開始しており5'非翻訳領域がないこと、そのアミノ酸配列がPTP-TD14のアミノ酸配列の途中240番目以降と高い相同性を示すことから、cKAL16は完全長のcDNAクローンではなく、完全長のcDNAはさらに5'末端側に延長しているものと考えられた。

そこで、cKAL16の塩基配列をもとにして、5'-RACE法〔Proc. Natl. Acad. Sci. USA, 85, 8998 (1988)〕により、さらに5'末端側に延長したcDNA断片を増幅し、クローン化した。cDNA断片の塩基配列を決定し、cKAL16の塩基配列とつなげたものを完全長のcDNAの塩基配列とし、配列番号1に示した。なお、この配列番号1の塩基配列とcKAL11の塩基配列を比較することにより、cKAL11の塩基配列の1～74番目もイントロン配列であることがわかった。配列番号1の塩基配列の64～4971番目には、1636アミノ酸からなるORFが存在しており、この領域をヒト新規PTPをコードする領域とし、そのアミノ酸配列を配列番号2に示した。このヒト新規PTPのアミノ酸配列を前述したラットPTP-TD14と比較すると、ラットPTP-TD14に比べてN末端が180アミノ酸長いが、図4に示すようにアミノ酸配列の181番目以降は部分的に挿入や欠失があるものの相同性を有していた。181番目以降部分のアミノ酸配列の相同性は83%であった。

該ヒト新規PTPのアミノ酸配列のC末端側1212～1454番目には、他のPTPと相同性をもつフォスファターゼ領域が存在した。ただし、第3図に示すように他のPTPの活性中心領域では保存されているアミノ酸配列〔Val His Cys Ser Ala Gly (Val/Ile) Gly Arg (Thr/Ser) Gly, J. Biol. Chem., 267, 140 (1992)〕内のアラニンがセリンに変換していた。また他のPTPでは保存されているが、該新規PTPでは保存されていないアミノ酸が14ヶ所存在した。また、770～1128番目には、ヒスチジンあるいはシステインを先頭に、プロリンに富んだ20～50アミノ酸からなる構造が15回繰り返す領域が存在した。該領域のヒスチジンおよびシステインに亜鉛などの金属イオンが配位するため、第5図に示すようなユニークな構造 (Zn-leaf様構造) をとると考えられる。本発明では、該繰り返し領域をヒスチジンドメイン (His Domain ; HD) と名付け、新規PTPをHD-PTPとした。

また、N末端の778アミノ酸は、第6図に示すように、酵母のMAPキナーゼ情報伝達経路に関与する蛋白質であるBR01〔Mol. Cell. Biol., 16, 2585 (1996)〕とも相同性を有する。さらに、BR01、BR01と相同性を有するラットPTP-TD14、

および線虫*C. elegans*の第3染色体の塩基配列解析からR10E12.1遺伝子 (GenBankアクセス番号 ; Z29561) にコードされていると推定される98kDa蛋白質 [Nature, 368, 32 (1994)] で共に保存されているアミノ酸配列モチーフ Lys Asp Asn Asp Phe Ile Tyr His Glu Xaa Val (Ser/Pro) が314~325番目に存在した。

また、配列番号2の719~730番目、745~750番目、898~905番目、950~957番目、1051~1058番目、1093~1102番目、1139~1145番目のHD付近には、SH3結合ドメインと考えられる配列が存在する。

また、配列番号2の414、665、679、922、924、998、990、1243番目のチロシン残基は、チロシンキナーゼによりリン酸化を受ける可能性のあるチロシン残基が存在する。

また、膜結合部位と考えられるような疎水性の高い領域がないため、細胞質に存在するタイプのフォスファターゼと考えられた。

以上の特徴を考え合わせると、HD-PTPは他の蛋白質と相互作用を行い、細胞内の情報伝達に関与しているPTPの可能性を有する。

また、C末端側には代謝回転の速い蛋白質に特異的に見出される配列である、プロリン、グルタミン酸、セリン、スレオニンからなるクラスターであるPEST配列 [Science, 234, 364 (1986)] が存在する。

実施例2 HD-PTP遺伝子の染色体位置の決定

配列番号3記載のHD-PTP遺伝子の584~604番目および2167~2186番目に相当する配列番号6および7に示すDNAをプライマーとして、ヒトゲノムDNAを鋳型にしてPCRを行ない、1.6kbのDNA断片が増幅することを確認した。PCRの温度条件は、94℃で4分間加熱後、1サイクルが94℃で1分間-64℃で4分間からなる反応を35サイクル繰り返し、最後に72℃で7分間加熱した。

Research Genetics社製のラディエーション・ハイブリッド・パネル (radiation hybrid panel) Gene-Bridge 4を購入し、93個のラディエーション・ハイブリッドクローンのDNA各25ngを鋳型にして、上記のPCRを行い、その結果をWhitehead

Institute/MIT Center for Genome ResearchのインターネットWebサイト (URL : <http://www-genome.wi.mit.edu/cgi-bin/contig/rhmapper.pl>) を利用して解析することにより、詳細な染色体マッピングを行った。

その結果、HD-PTP遺伝子はWhitehead 研究所の染色体上の位置を示すマーカーWI-11814とWI-16174の間にあることが判明した。この位置は、第3染色体短腕 (3p) の3p21.3に存在するマーカーD3S3888 (テロメア側) とD3S3334 (セントロメア側) の間であった。3p21.3という染色体位置は、肺癌をはじめ、頸部癌や大腸癌で高頻度にLOHがみられる位置と一致していた。またこの位置は、SCIDマウスで腫瘍性を示すような、ヒト第3染色体およびマウスA9細胞の雑種融合細胞で共通して欠失している、ヒト第3染色体上の1.6cM (センチモルガン) の領域 [Genes, Chromosome & Cancer, 20, 329 (1997)] 内にあり、まだ見出されていない3p21.3に存在する癌抑制遺伝子と考えられた。

実施例3 HD-PTP遺伝子のゲノムDNAの塩基配列

実施例1で得られたcKAL11は、イントロンが残存しているcDNAクローンであるので、大部分はゲノムDNAの塩基配列と一致すると考えられたが、一部のイントロンがスプライスされている可能性もある。そこで、ヒトゲノムDNAを鋳型にして、cKAL11の各エクソンと考えられた領域を含む断片をPCRによって増幅し、その塩基配列をcKAL11の塩基配列と比較することにより、そのエクソン内にイントロンが存在しないかどうかを確認した。ただし、配列番号5の塩基配列1741～3679番目に相当するエクソンは長いので、6分割して増幅することにした。

配列番号8～37に記載した塩基配列を有する15組のプライマーセット (配列番号8と9、配列番号10と11、配列番号12と13、配列番号14と15、配列番号16と17、配列番号18と19、配列番号20と21、配列番号22と23、配列番号24と25、配列番号26と27、配列番号28と29、配列番号30と31、配列番号32と33、配列番号34と35、配列番号36と37) を用いてPCRを行った結果、配列番号4の塩基配列1348～1654番目に相当するエクソン中に99bpからなる1イントロン、塩基配列

1741～3679番目に相当するエクソン中に93bpからなる1イントロンがさらに存在することがわかった。また、上述の方法により得られたHD-PTPゲノムDNAよりもさらに5'側のHD-PTPゲノムDNAの塩基配列を、実施例1で得られたHD-PTPのcDNAの塩基配列を利用したプライマー・ウォーキングにより解析した。その結果、配列番号1の塩基配列を有するcDNAは25エクソンからなり、エクソン間に24イントロンが挿入されていることが明らかとなった。第1エクソンと第2エクソンの間の第1イントロンは約5.5kb、また第2エクソンと第3エクソンの間の第2イントロンは約8kbと長大なものであり、第1エクソンから第25エクソンまで約22kbに渡っていた。この2つの長いイントロンについては、エクソンと隣接する領域の塩基配列のみ決定した。HD-PTP遺伝子のゲノムDNAのうち、配列番号40に第1エクソンおよび隣接する第1イントロンの5'端の領域、配列番号41に第2エクソンおよび隣接する第1イントロンの3'端と第2イントロンの5'端の領域、配列番号3に第2イントロンの3'端とそれに隣接する第3エクソンから第25エクソンまでの領域の塩基配列をそれぞれ示した。配列番号40の1～131 (エクソン1)、配列番号41の404～478 (エクソン2)、配列番号42の529～656 (エクソン3)、881～957 (エクソン4)、1625～1674 (エクソン5)、1791～1922 (エクソン6)、2201～2281 (エクソン7)、2357～2488 (エクソン8)、2579～2626 (エクソン9)、3006～3062 (エクソン10)、3185～3243 (エクソン11)、3381～3460 (エクソン12)、3573～3687 (エクソン13)、3766～3831 (エクソン14)、4221～4366 (エクソン15)、4652～4963 (エクソン16)、5039～5193 (エクソン17)、5293～5444 (エクソン18)、5531～5710 (エクソン19)、5804～7562 (エクソン20)、7657～7841 (エクソン21)、7968～8072 (エクソン22)、8157～8295 (エクソン23)、8388～8501 (エクソン24) および8580～9307番目 (エクソン25) がエクソンに当たる領域である。

実施例4 肺癌細胞でのHD-PTP遺伝子の変異の検出

(1) 肺癌患者における染色体3p21.3のLOH

文献〔Cancer Res., 57, 1344 (1997)〕の方法に基づいて、PCRによりヒト染色体の3p21.3に存在するマイクロサテライトマーカーであるD3S3564、D3S3559、D3S3582、D3S1568についてマイクロサテライト多型を解析することにより、3p21.3付近のLOHの有無について、肺癌患者の癌組織と正常組織30組を調べた。

その結果、調べた癌組織の約40%において1つ以上のマーカーの欠失がみられた。

(2) 肺小細胞癌セルラインにおけるHD-PTP遺伝子の変異の検出

SSCP (single strand conformational polymorphism) 法〔Proc. Natl. Acad. Sci., 86, 2766 (1989)〕により、肺小細胞癌のゲノムDNA中のHD-PTP遺伝子の塩基配列の変異の検出を試みた。

実施例3で用いたエクソン増幅用のPCRプライマーを用いて、肺小細胞癌系セルライン〔Lu-130 (理化学研究所細胞銀行 RCB0465)、Lu-135 (理化学研究所細胞銀行 RCB0468)、NCI-H69 (ATCC HTB-119)、NCI-H82 (ATCC HTB-175)、NCI-N417 (ATCC CRL-5809)、RERF-LC-MA (JCRB/HSRRB細胞バンク JCRB0812)、SBC-5 (JCRB/HSRRB細胞バンク JCRB0819)、NCI-H526 (ATCC CRL-5811)、NCI-H209 (ATCC HTB-172)、NCI-H841 (ATCC CRL-5845)、NCI-H774 (ATCC CRL-5842)、MS-18〕のゲノムDNAを鋳型として用い、PCRを行った。PCR後、得られた反応サンプルを用い、非変性条件下でポリアクリルアミドゲル電気泳動を行った。泳動後、ゲルを銀染色した。

正常なHD-PTP遺伝子を用いた場合と比較し、移動度の変化したPCR増幅DNA断片をゲルから抽出した。該DNAを鋳型として用い、再度、上記操作を繰り返し、増幅DNA断片を精製した。

該DNA断片を鋳型として用い、PCRプライマーを利用した塩基配列の決定を行った。

その結果、RERF-LC-MAにおいては、配列番号3の塩基番号6969番目、配列番号1の塩基番号3358番目に相当する塩基がCからTに変異していた。この変異に

より、コードするアミノ酸残基がプロリンからセリンに置換される。該プロリン残基は、SH3結合モチーフ部位中に存在するため、セリンへの変異により、HD-PTPが正常に機能しない可能性がある。

また、癌細胞系セルラインあるいは患者癌組織のサンプル（合計総数325）について、さらに同様の解析を行なったところ、6個のサンプル（全体の1.8%に相当）において、配列番号3の塩基番号7724番目、配列番号1の塩基番号4019番目に相当する塩基がCからTに変異していた。この変異により、コードするアミノ酸残基がアラニンからバリンに置換される。

実施例5 癌細胞におけるHD-PTP遺伝子の発現量の変動

9種類のヒト癌細胞系セルライン〔胃癌；MKN1 (RCB1003)、MKN28 (RCB1000)、MKN45、MKN74 (RCB1002)、KATO III (ATCC HTB-103)、腎癌；KPK1 [Journal of Urology, 128, 1117 (1982)]、KPK13 [Journal of Urology, 128, 1117 (1982)]、口腔癌；KB (ATCC CCL-17)、リンパ種；KAL-1〕および正常繊維芽細胞セルラインWI-38 (ATCC CCL-75) から全RNAあるいはpoly A RNAを常法により抽出した。全RNAを20 μ gまたはpoly A RNAを2 μ g用いて、実施例1で得られたcKAL16をプローブにしてノーザンブロットを行なった。2種類の胃癌細胞系セルラインMKN74およびKATO IIIでは、全くバンドがみられず、HD-PTP遺伝子が発現していないと考えられた。

実施例6 各組織でのHD-PTP遺伝子の発現

オス成体ラットの脳、肺、心臓、耳下腺、胃、小腸、大腸、肝臓、膵臓、腎臓、膀胱、精巣、精巣上体、精囊、脾臓、胸腺、リンパ球、甲状腺および副腎からそれぞれ全RNAを常法により抽出し、実施例5に記載の方法に準じて、実施例1で得られたcKAL16をプローブとして用い、ノーザン・ブロット・ハイブリダイゼーションを行なった。

ハイブリダイゼーションの結果、RNAの分解が激しかった膵臓と脾臓を除く全ての臓器で、ラットHD-PTP mRNAと考えられるバンドが検出されたため、HD-PTP

遺伝子は全ての臓器で普遍的に発現していると考えられた。

HD-PTPは細胞内フォスファターゼであり、全ての臓器に普遍的に存在すると考えられるため、HD-PTPは細胞内での情報伝達に重要な役割を果たしている分子である可能性が高い。

実施例7 遺伝子組換え法による動物細胞でのHD-PTPの発現

ハイグロマイシン耐性遺伝子およびEBウイルス複製起点を有するベクターEB0-pcD [Mol. Cell. Biol., 8, 2837 (1988)] に、ヒトT細胞白血病ウイルスIのLTR由来のプロモーターであるSR α プロモーター [Mol. Cell. Biol., 8, 466 (1988)] を組み込んだプラスミドpCD-EB (九州大学 早川浩博士より供与) のSR α プロモーターの直後のXhoI/BamHIサイトにマルチクローニングサイト (XhoI-NotI-XbaI-KpnI-BamHI) を挿入し、動物細胞発現用ベクターpCDL81を作製した。マルチクローニングサイトは、配列番号38および39に示した塩基配列をDNA合成機で合成し、作製した。pCDL81に実施例1でクローン化したcKAL16のHD-PTP cDNAを挿入し、HD-PTP発現プラスミドpDKL4fを作製した。なお、HD-PTP発現プラスミドpDKL4fを含有する形質転換体*Escherichia coli* DH5 α /pDKL4fは、平成10年9月11日付けで工業技術院生命工学工業技術研究所 (日本国茨城県つくば市東1丁目1番3号 郵便番号305-8556) にFERM BP-6499として寄託されている。

リン酸カルシウム法を用いて、マウス細胞株NIH3T3あるいはラット細胞株Bat-2に発現プラスミドDNAを導入し、ハイグロマイシン耐性により形質転換細胞を選択した。選択された形質転換細胞から常法によりDNAを抽出し、アガロースゲル電気泳動を行い、ゲルをフィルターに転写後、cKAL16 cDNAをプローブとしてサザン・ブロット・ハイブリダイゼーションを行い、形質転換細胞に導入したヒトHD-PTP cDNAが保持されていることを確認した。また形質転換細胞から常法によりRNAを抽出し、実施例4と同様にしてノーザン・ブロット・ハイブリダイゼーションを行い、導入したヒトHD-PTP遺伝子の発現を検出した。

ヒトHD-PTP遺伝子の発現の認められた形質転換細胞は、実施例1でみられたような軟寒天中でのコロニー形成を示すようになった。このコロニー形成能はノーザンブロットによってHD-PTP遺伝子の発現量が高いものほど強い傾向にあった。この形質転換細胞 2×10^7 個を、放射線照射した胸腺欠損ヌードマウスの皮下に注入し、腫瘍形成するかどうか60日間観察したが、注入後腫瘍を形成したマウスはいなかった。したがってこのコロニー形成能は腫瘍形成能とは無関係の性質であると考えられた。

これらの形質転換細胞の細胞抽出液を用いて、文献〔J. Biol. Chem., 269, 13614 (1994)〕の方法に基づいた、p-ニトロフェノールリン酸を基質としてフォスファターゼ活性の測定を行った。形質転換していない細胞の細胞抽出液と比較してHD-PTP遺伝子で形質転換した細胞のフォスファターゼ活性は約2倍に上昇していた。

実施例8 HD-PTPの細胞内の局在

GFP (Green fluorescent protein) との融合蛋白発現用ベクターpEGFP-N1 (Clontech社) のAccIサイトにHD-PTP cDNAを挿入し、HD-PTPのC末端側にGFPを融合させた蛋白質 (HD-PTP/GFP融合蛋白質) を発現させるプラスミドを作製した。該プラスミドをヒト細胞株293tsA1609neo〔Mol. Cell. Biol., 7, 379 (1987)〕に導入し、HD-PTP/GFP融合蛋白質を発現させた。形質転換細胞を蛍光顕微鏡で観察し、GFPの蛍光を検出することにより、HD-PTP/GFP融合蛋白質の細胞内の局在を調べた。

該検出の結果、HD-PTP/GFP融合蛋白質は、細胞膜、特に核近傍の細胞質に複数の塊状になって存在しており、細胞膜や核には存在しないことがわかった。このことは特定の細胞内器官に局在する可能性を示唆した。

HD-PTPのアミノ酸配列から推定した分子量は約150kDaであり、SDS-PAGEによる分子量測定でも、HD-PTP/GFP融合蛋白質は、予想される分子量に近い150kDaの位置に検出された。該融合蛋白質は、非変性条件のPAGEでは70kDaの位置に検

出されたことより、HD-PTPは天然の非変性状態では非常にコンパクトな球状の構造を有すると推察された。

産業上の利用可能性

本発明の新規チロシンフォスファターゼおよびその遺伝子を用いることにより、癌の診断および治療が可能となる。

「配列表フリーテキスト」

配列番号6－人工配列の説明：HD-PTP遺伝子増幅のためのセンスプライマー

配列番号7－人工配列の説明：HD-PTP遺伝子増幅のためのアンチセンスプライマー

配列番号8－人工配列の説明：エクソン14を含むHD-PTP遺伝子の配列番号3における3692-3939位増幅のためのセンスプライマー

配列番号9－人工配列の説明：エクソン14を含むHD-PTP遺伝子の配列番号3における3692-3939位増幅のためのアンチセンスプライマー

配列番号10－人工配列の説明：エクソン15を含むHD-PTP遺伝子の配列番号3における4159-4421位増幅のためのセンスプライマー

配列番号11－人工配列の説明：エクソン15を含むHD-PTP遺伝子の配列番号3における4159-4421位増幅のためのアンチセンスプライマー

配列番号12－人工配列の説明：エクソン16を含むHD-PTP遺伝子の配列番号3における4625-5002位増幅のためのセンスプライマー

配列番号13－人工配列の説明：エクソン16を含むHD-PTP遺伝子の配列番号3における4625-5002位増幅のためのアンチセンスプライマー

配列番号14－人工配列の説明：エクソン17および18を含むHD-PTP遺伝子の配列番号3における5014-5473位増幅のためのセンスプライマー

配列番号15－人工配列の説明：エクソン17および18を含むHD-PTP遺伝子の配列番号3における5014-5473位増幅のためのアンチセンスプライマー

- 配列番号16－人工配列の説明：エクソン19および20の一部を含むHD-PTP遺伝子の配列番号3における5499-5984位増幅のためのセンスプライマー
- 配列番号17－人工配列の説明：エクソン19および20の一部を含むHD-PTP遺伝子の配列番号3における5499-5984位増幅のためのアンチセンスプライマー
- 配列番号18－人工配列の説明：エクソン20の一部を含むHD-PTP遺伝子の配列番号3における5943-6281位増幅のためのセンスプライマー
- 配列番号19－人工配列の説明：エクソン20の一部を含むHD-PTP遺伝子の配列番号3における5943-6281位増幅のためのアンチセンスプライマー
- 配列番号20－人工配列の説明：エクソン20の一部を含むHD-PTP遺伝子の配列番号3における6191-6579位増幅のためのセンスプライマー
- 配列番号21－人工配列の説明：エクソン20の一部を含むHD-PTP遺伝子の配列番号3における6191-6579位増幅のためのアンチセンスプライマー
- 配列番号22－人工配列の説明：エクソン20の一部を含むHD-PTP遺伝子の配列番号3における6478-6908位増幅のためのセンスプライマー
- 配列番号23－人工配列の説明：エクソン20の一部を含むHD-PTP遺伝子の配列番号3における6478-6908位増幅のためのアンチセンスプライマー
- 配列番号24－人工配列の説明：エクソン20の一部を含むHD-PTP遺伝子の配列番号3における6866-7290位増幅のためのセンスプライマー
- 配列番号25－人工配列の説明：エクソン20の一部を含むHD-PTP遺伝子の配列番号3における6866-7290位増幅のためのアンチセンスプライマー
- 配列番号26－人工配列の説明：エクソン20の一部を含むHD-PTP遺伝子の配列番号3における7244-7639位増幅のためのセンスプライマー
- 配列番号27－人工配列の説明：エクソン20の一部を含むHD-PTP遺伝子の配列番号3における7244-7639位増幅のためのアンチセンスプライマー
- 配列番号28－人工配列の説明：エクソン21を含むHD-PTP遺伝子の配列番号3における7628-7888位増幅のためのセンスプライマー

配列番号29－人工配列の説明：エクソン21を含むHD-PTP遺伝子の配列番号3における7628-7888位増幅のためのアンチセンスプライマー

配列番号30－人工配列の説明：エクソン22を含むHD-PTP遺伝子の配列番号3における7897-8131位増幅のためのセンスプライマー

配列番号31－人工配列の説明：エクソン22を含むHD-PTP遺伝子の配列番号3における7897-8131位増幅のためのアンチセンスプライマー

配列番号32－人工配列の説明：エクソン23を含むHD-PTP遺伝子の配列番号3における8091-8379位増幅のためのセンスプライマー

配列番号33－人工配列の説明：エクソン23を含むHD-PTP遺伝子の配列番号3における8091-8379位増幅のためのアンチセンスプライマー

配列番号34－人工配列の説明：エクソン24を含むHD-PTP遺伝子の配列番号3における8317-8562位増幅のためのセンスプライマー

配列番号35－人工配列の説明：エクソン24を含むHD-PTP遺伝子の配列番号3における8317-8562位増幅のためのアンチセンスプライマー

配列番号36－人工配列の説明：エクソン25の一部を含むHD-PTP遺伝子の配列番号3における8492-8905位増幅のためのセンスプライマー

配列番号37－人工配列の説明：エクソン25の一部を含むHD-PTP遺伝子の配列番号3における8492-8905位増幅のためのアンチセンスプライマー

配列番号38－人工配列の説明：マルチクローニングサイト (XhoI NotI XbaI KpnI BamHI) リンカーのための合成DNA

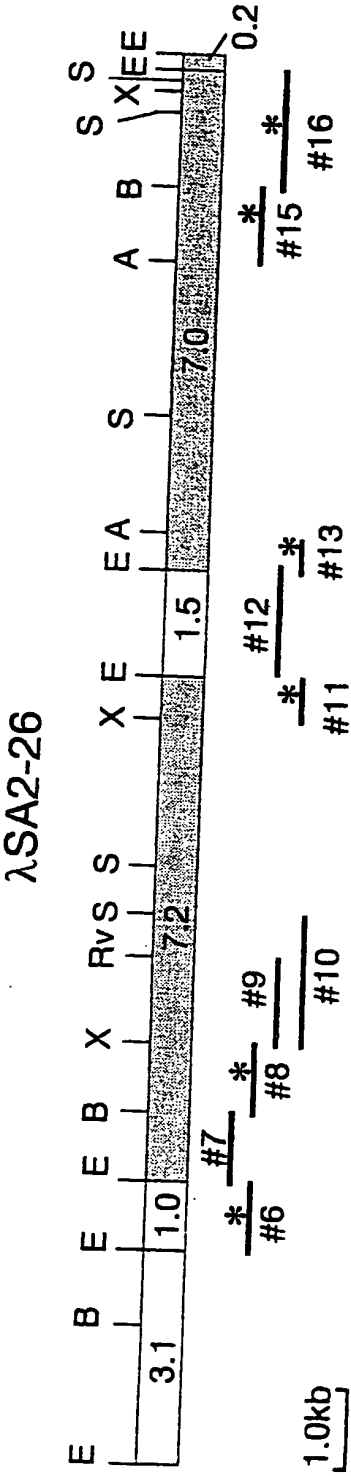
配列番号39－人工配列の説明：マルチクローニングサイト (XhoI NotI XbaI KpnI BamHI) リンカーのための合成DNA

請求の範囲

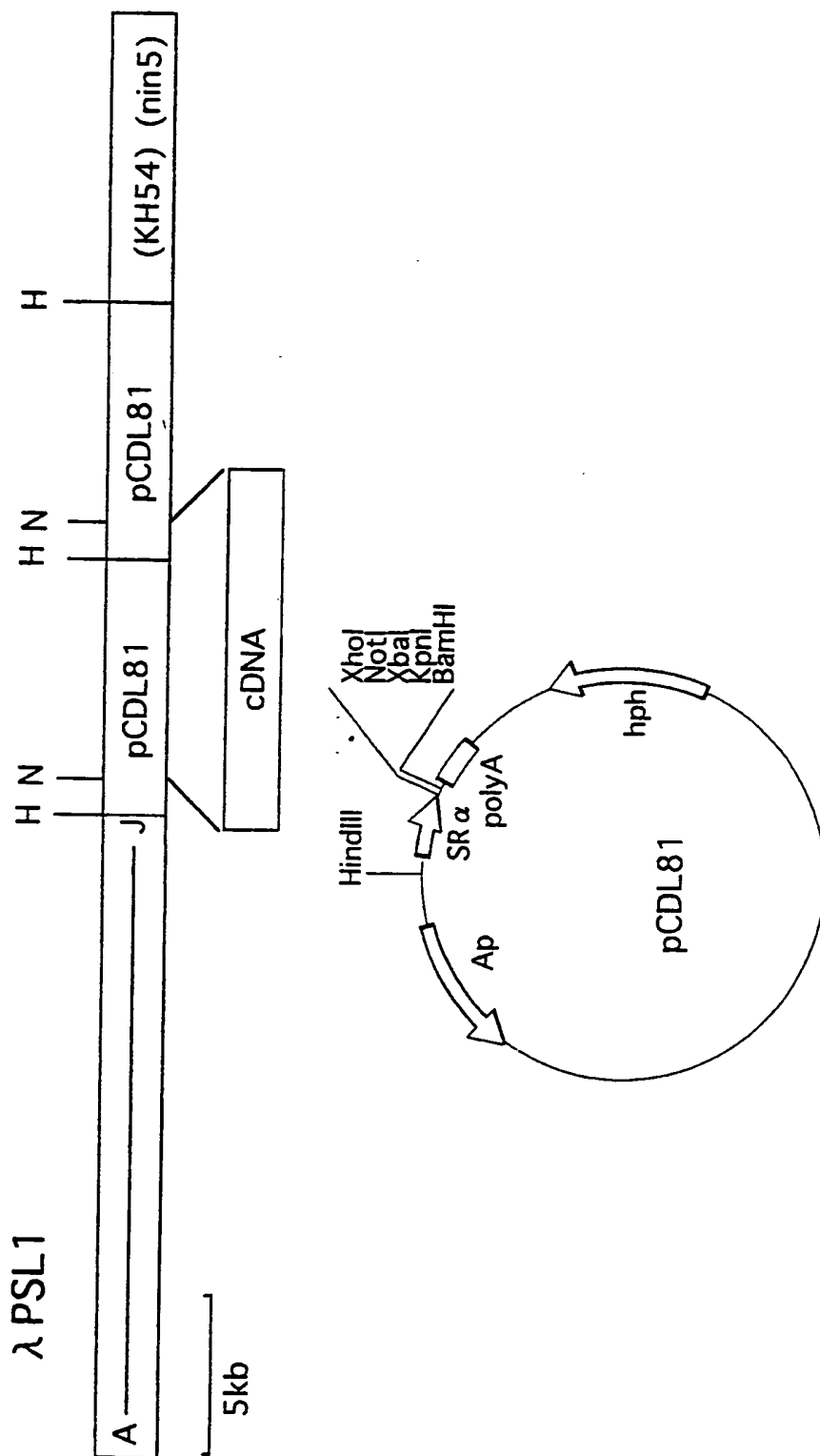
1. 配列番号 2 記載のアミノ酸配列からなる蛋白質。
2. 配列番号 2 記載の蛋白質の有するアミノ酸配列において 1 もしくは数個のアミノ酸が欠失、置換もしくは付加されたアミノ酸配列からなり、かつチロシンフォスファターゼ活性を有する蛋白質。
3. 請求項 1 または 2 記載の蛋白質をコードする DNA。
4. 配列番号 1、3、40 または 41 記載の塩基配列を有する DNA。
5. 請求項 3 または 4 記載の DNA とストリンジェントな条件下でハイブリダイズし、かつチロシンフォスファターゼ活性を有する蛋白質をコードする DNA。
6. 請求項 3～5 のいずれか 1 項に記載の DNA とベクター DNA とを含有する組換え体 DNA。
7. 請求項 6 記載の組換え体 DNA を宿主細胞に導入して得られる形質転換体。
8. 請求項 7 記載の形質転換体を培地に培養し、培養物中に請求項 1 または 2 記載の蛋白質を生成蓄積させ、該培養物から該蛋白質を採取することを特徴とする請求項 1 または 2 記載の蛋白質の製造方法。
9. 請求項 1 または 2 記載の蛋白質を有効成分として含む、癌の治療薬。
10. 請求項 3～5 のいずれか 1 項に記載の DNA を有効成分とする癌の治療薬。
11. 請求項 3～5 のいずれか 1 項に記載の DNA を含有する、癌の遺伝子治療用ベクター。
12. 請求項 3～5 のいずれか 1 項に記載の DNA の塩基配列のうち、連続した 10～60 残基の塩基配列を有するオリゴヌクレオチド、または該オリゴヌクレオチドと相補的な配列を有するオリゴヌクレオチド、およびこれらオリゴヌクレオチドの誘導体であるオリゴヌクレオチド誘導体。
13. 配列番号 24、25、28 または 29 で表される塩基配列を有する請求項 12 記載のオリゴヌクレオチド。

14. 請求項 1 2 または 1 3 記載のオリゴヌクレオチドを用いた、癌の診断方法。
15. 請求項 1 2 または 1 3 記載のオリゴヌクレオチドを含有する、癌の診断薬。
16. 請求項 1 または 2 記載の蛋白質を認識する抗体。
17. 請求項 1 6 記載の抗体を用いて、請求項 1 または 2 記載の蛋白質を免疫学的に検出する方法。
18. 請求項 1 6 記載の抗体を用いて、請求項 1 または 2 記載の蛋白質を免疫学的に定量する方法。
19. 請求項 1 6 記載の抗体を用いる、癌の判定方法。
20. 請求項 1 6 記載の抗体を有効成分とする、癌の診断薬。

第1図



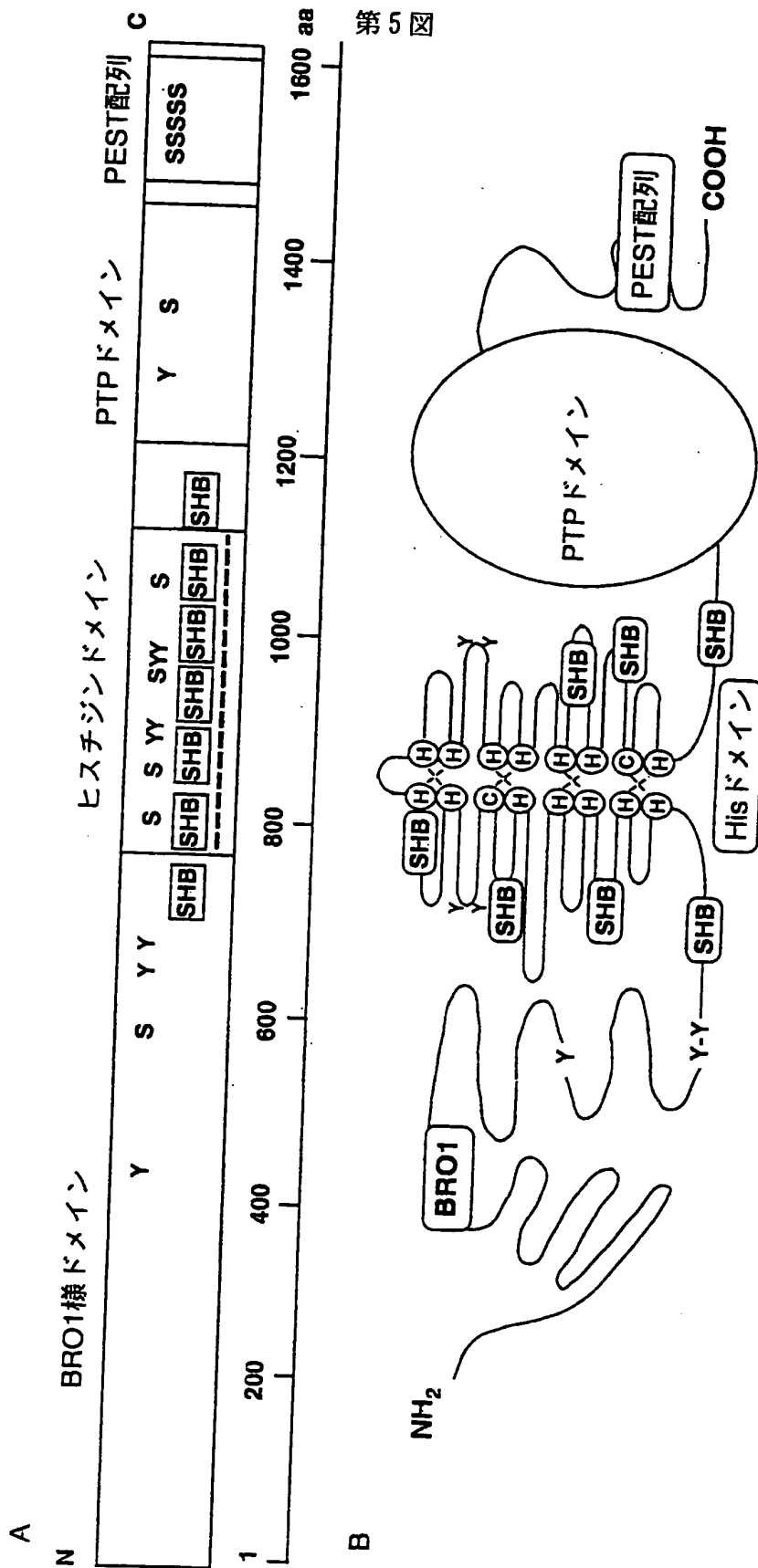
第2図



第3図

[illegible]

[illegible]



1 MEAVRMPVWLDLKEAGDFHFQPAVKKFLVKNYGNP--EAYNE-ELKKLELLRQNA--VRVPRDFEGCSV-LRKYLQQLHY-LQSRVPMGSGQEAAPVTVTTEIFSGKSAVAHEDIKYE
 1 MKPYLFDLKLKDKTEKLDWKGLSSYLKKSYGSSQWRIFYDEKATSELDHLRNVANGELAPSSISEQNLYYSFLEHLYFRLGSGSRLKMDFTWYDAEYSSAQKGLKYTQHTLAFE
 114 QACILYNLGAHSMLGAMDKRVSEEGMKVSCTHFQCAAGAFAYLREHFPAQYSVDMRSRQILTLNVNMLGQAQECLEKSMLD--NRKSF-LVARISQVVDYKYACRALENPDITASLL
 117 KSCTLFNIAVI---FTQIARENINEDYKNSIANLTKAFSCFEYLSENFLNSPSVDLQSENTRFLANICHAEAQELFVLKLLNDQISSKQYTLISKLSRATCNLF-QKCHDFMKEIDDDVA
 231 GRIQKDWKKLVQMKIYYFAAVALHMGKQAEQKFGERVAYFQSALDKPNEAIKLAKGQPDVQ-DALRFTMDVIGGKYNsAKKDNDFIYHEAVPA-----LDTL-PVKGAPLVKPLP
 233 IYGEPKWKTVTCKLHFYKSL SAYYHGLHLEENRVGEAIAFLDFSMQQLISSLPFKTWLVEFIDFGFKETLE---KKQKELIKDNDFIYHESVPAVVQVDSIKALDAIKSPTWEKILE
 343 VNPTDPAVTGPDIFAKLPMAAHEASSLYSEEKALLREMMAKIEDKNEVLDOFMSMQLDPETVDNLDAYSHIPQLMKCAALSVRPDVTVRNLVQSMQVLSGVFTDVEASLKDIROLL
 350 PYMQDVANKYDSL YRGIIPLDVYEKESIYSEEKATLLRKQVEETETANLEYSSSFIETNL-PRLLSDLEKQFSDGNIFSNITDQQLMRQIQTWCKFIQ--TNEFRDIEEQMVKI--VF
 463 EDELELEQKFQEA VQAGAGAI SITSKAELAEVRREWAKYMEVHEKASFTNSELHRAMNLHVGNLRLLSGPLDQVRAALPTALSPEDK-AVLQNLKRILAKVQEMRQORVSLEQQLREL IQ
 465 KRKQILEILSALPNDQKENVT-KLKSSLVAASNSDEKLFAC-VKPHIVEINLLNDNGKIWKKFEDEFNRNTPPPQSLDIDDTKNDKILELLKQVKGHAEDLRTLKEERSNLSLROEIN
 582 KODITASL-VTTDHS--EMKKLFEELKKYDQLKVYLEQNLAQDRVL-----CALTE--ANVQYAAVRRVLSDLQKWNSTLQTLVASYEAYEDLMKKSQEGRDFYADLESKVAALLERT
 583 NDDITKLLIINKGKSDVELKDLFEVELEKFEPLSTRIEATIIYKQSSMIDDIIKAKLDEIFHL SNFKDKSSGEEKFLEDRKNFFDKLQEA VKSFSIFASDLPKGIEFYDSL FNMSRDLAERV
 693 QSTQAREARQQLLDRELKKKPPRPTAPKPLLPREEE-EAVEAGDPPEELRSLPPDMVAGPRLPDTFLGSATPLHFPSPFPSS 778
 703 RVAKQTEDSTANSPAPPLPPLDSKASVVGPPPLLPQKSAAFQSLSRQGLNLGDQFQNLKISAGSDLPQG--PGIPPRTYEASPYAAT 787

図 6
 將

35	40	45	
ctg gag ttg ctc aga cag aat gct gtc cgt gtc cca cga gac ttt gag			252
Leu Glu Leu Leu Arg Gln Asn Ala Val Arg Val Pro Arg Asp Phe Glu			
50	55	60	
ggc tgt agt gtc ctc cgc aag tac ctc ggc cag ctt cat tac ctg cag			300
Gly Cys Ser Val Leu Arg Lys Tyr Leu Gly Gln Leu His Tyr Leu Gln			
65	70	75	
agt cgg gtc ccc atg ggc tgc ggc cag gag gcc gct gtc cct gtc acc			348
Ser Arg Val Pro Met Gly Ser Gly Gln Glu Ala Ala Val Pro Val Thr			
80	85	90	95
tgg aca gag atc ttc tca ggc aag tct gtg gcc cat gag gac atc aag			396
Trp Thr Glu Ile Phe Ser Gly Lys Ser Val Ala His Glu Asp Ile Lys			
100	105	110	
tac gag cag gcc tgt att ctc tac aac ctt gga gcg ctg cac tcc atg			444
Tyr Glu Gln Ala Cys Ile Leu Tyr Asn Leu Gly Ala Leu His Ser Met			
115	120	125	
ctg ggg gcc atg gac aag cgg gtg tct gag gag ggc atg aag gtc tcc			492
Leu Gly Ala Met Asp Lys Arg Val Ser Glu Glu Gly Met Lys Val Ser			
130	135	140	
tgt acc cat ttc cag tgc gca gcc ggc gcc ttc gcc tac cta cgg gag			540
Cys Thr His Phe Gln Cys Ala Ala Gly Ala Phe Ala Tyr Leu Arg Glu			
145	150	155	
cac ttc cct caa gcc tac agc gtc gac atg agc cgc cag atc ctt acg			588
His Phe Pro Gln Ala Tyr Ser Val Asp Met Ser Arg Gln Ile Leu Thr			
160	165	170	175
ctc aac gtc aac ctc atg ctg ggc cag gct cag gag tgc ctc ctg gag			636
Leu Asn Val Asn Leu Met Leu Gly Gln Ala Gln Glu Cys Leu Leu Glu			
180	185	190	
aag tcg atg ttg gac aac agg aag agc ttt ctg gtg gcc cgc atc agt			684
Lys Ser Met Leu Asp Asn Arg Lys Ser Phe Leu Val Ala Arg Ile Ser			
195	200	205	

配 列 表

SEQUENCE LISTING

<110> KYOWA HAKKO KOGYO CO., LTD

<120> Novel tyrosine phosphatase

<130> 11201WO1

<140>

<141>

<150> JP 99/108842

<151> 1999-4-16

<160> 39

<170> PatentIn Ver. 2.0

<210> 1

<211> 5234

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (64)..(4971)

<400> 1

tggctgagcc agcagctgca gcagctgcgg gaggggccgg gtggccggcg ggtgccaccc 60
 gcc atg gag gcc gtg ccc cgc atg ccc gtg atc tgg ctg gac ctg aag 108
 Met Glu Ala Val Pro Arg Met Pro Val Ile Trp Leu Asp Leu Lys
 1 5 10 15

gag gcc ggt gac ttt cac ttc cag cca gct gtg aag aag ttt gtc ctg 156
 Glu Ala Gly Asp Phe His Phe Gln Pro Ala Val Lys Lys Phe Val Leu
 20 25 30

aag aat tat gga gag aac cca gaa gcc tac aat gaa gaa ctg aag aag 204
 Lys Asn Tyr Gly Glu Asn Pro Glu Ala Tyr Asn Glu Glu Leu Lys Lys

35	40	45	
ctg gag ttg ctc aga cag aat gct gtc cgt gtc cca cga gac ttt gag			252
Leu Glu Leu Leu Arg Gln Asn Ala Val Arg Val Pro Arg Asp Phe Glu			
50	55	60	
ggc tgt agt gtc ctc cgc aag tac ctc ggc cag ctt cat tac ctg cag			300
Gly Cys Ser Val Leu Arg Lys Tyr Leu Gly Gln Leu His Tyr Leu Gln			
65	70	75	
agt cgg gtc ccc atg ggc tcg ggc cag gag gcc gct gtc cct gtc acc			348
Ser Arg Val Pro Met Gly Ser Gly Gln Glu Ala Ala Val Pro Val Thr			
80	85	90	95
tgg aca gag atc ttc tca ggc aag tct gtg gcc cat gag gac atc aag			396
Trp Thr Glu Ile Phe Ser Gly Lys Ser Val Ala His Glu Asp Ile Lys			
100	105	110	
tac gag cag gcc tgt att ctc tac aac ctt gga gcg ctg cac tcc atg			444
Tyr Glu Gln Ala Cys Ile Leu Tyr Asn Leu Gly Ala Leu His Ser Met			
115	120	125	
ctg ggg gcc atg gac aag cgg gtg tct gag gag ggc atg aag gtc tcc			492
Leu Gly Ala Met Asp Lys Arg Val Ser Glu Glu Gly Met Lys Val Ser			
130	135	140	
tgt acc cat ttc cag tgc gca gcc ggc gcc ttc gcc tac cta cgg gag			540
Cys Thr His Phe Gln Cys Ala Ala Gly Ala Phe Ala Tyr Leu Arg Glu			
145	150	155	
cac ttc cct caa gcc tac agc gtc gac atg agc cgc cag atc ctt acg			588
His Phe Pro Gln Ala Tyr Ser Val Asp Met Ser Arg Gln Ile Leu Thr			
160	165	170	175
ctc aac gtc aac ctc atg ctg ggc cag gct cag gag tgc ctc ctg gag			636
Leu Asn Val Asn Leu Met Leu Gly Gln Ala Gln Glu Cys Leu Leu Glu			
180	185	190	
aag tcg atg ttg gac aac agg aag agc ttt ctg gtg gcc cgc atc agt			684
Lys Ser Met Leu Asp Asn Arg Lys Ser Phe Leu Val Ala Arg Ile Ser			
195	200	205	

gca cag gtg gta gat tac tac aag gag gca tgc cgg gcc ttg gag aac 732
 Ala Gln Val Val Asp Tyr Tyr Lys Glu Ala Cys Arg Ala Leu Glu Asn
 210 215 220

ccc gac act gcc tca ctg ctg ggc cgg atc cag aag gac tgg aag aaa 780
 Pro Asp Thr Ala Ser Leu Leu Gly Arg Ile Gln Lys Asp Trp Lys Lys
 225 230 235

ctt gtg cag atg aag atc tac tac ttc gca gcc gtg gct cat ctg cac 828
 Leu Val Gln Met Lys Ile Tyr Tyr Phe Ala Ala Val Ala His Leu His
 240 245 250 255

atg gga aag cag gcc gag gag cag cag aag ttc ggg gag cgg gtt gca 876
 Met Gly Lys Gln Ala Glu Glu Gln Gln Lys Phe Gly Glu Arg Val Ala
 260 265 270

tac ttc cag agc gcc ctg gac aag ccc aat gaa gcc atc aag ttg gcc 924
 Tyr Phe Gln Ser Ala Leu Asp Lys Pro Asn Glu Ala Ile Lys Leu Ala
 275 280 285

aag ggc cag cct gac act gtg caa gac gcg ctt cgc ttc act atg gat 972
 Lys Gly Gln Pro Asp Thr Val Gln Asp Ala Leu Arg Phe Thr Met Asp
 290 295 300

gtc att ggg gga aag tac aat tct gcc aag aag gac aac gac ttc att 1020
 Val Ile Gly Gly Lys Tyr Asn Ser Ala Lys Lys Asp Asn Asp Phe Ile
 305 310 315

tac cat gag gct gtc cca gca ttg gac acc ctt cag cct gta aaa gga 1068
 Tyr His Glu Ala Val Pro Ala Leu Asp Thr Leu Gln Pro Val Lys Gly
 320 325 330 335

gcc ccc ttg gtg aag ccc ttg cca gtg aac ccc aca gac cca gct gtt 1116
 Ala Pro Leu Val Lys Pro Leu Pro Val Asn Pro Thr Asp Pro Ala Val
 340 345 350

aca ggc cct gac atc ttt gcc aaa ctg gta ccc atg gct gcc cac gag 1164
 Thr Gly Pro Asp Ile Phe Ala Lys Leu Val Pro Met Ala Ala His Glu
 355 360 365

gcc tcg tca ctg tac agt gag gag aag gcc aag ctg ctc cgg gag atg 1212
 Ala Ser Ser Leu Tyr Ser Glu Glu Lys Ala Lys Leu Leu Arg Glu Met

370	375	380	
atg gcc aag att gag gac aag aat gag gtc ctg gac cag ttc atg gat Met Ala Lys Ile Glu Asp Lys Asn Glu Val Leu Asp Gln Phe Met Asp 385 390 395	1260		
tca atg cag ttg gat ccc gag acg gtg gac aac ctt gat gcc tac agc Ser Met Gln Leu Asp Pro Glu Thr Val Asp Asn Leu Asp Ala Tyr Ser 400 405 410 415	1308		
cac atc cca ccc cag ctc atg gag aag tgc gcg gct ctc agc gtc cgg His Ile Pro Pro Gln Leu Met Glu Lys Cys Ala Ala Leu Ser Val Arg 420 425 430	1356		
ccc gac act gtc agg aac ctt gta cag tcc atg caa gtg ctg tca ggt Pro Asp Thr Val Arg Asn Leu Val Gln Ser Met Gln Val Leu Ser Gly 435 440 445	1404		
gtg ttc acg gat gtg gag gct tcc ctg aag gac atc aga gat ctg ttg Val Phe Thr Asp Val Glu Ala Ser Leu Lys Asp Ile Arg Asp Leu Leu 450 455 460	1452		
gag gag gat gag ctg cta gag cag aag ttt cag gag gcg gtg ggc cag Glu Glu Asp Glu Leu Leu Glu Gln Lys Phe Gln Glu Ala Val Gly Gln 465 470 475	1500		
gca ggg gcc atc tcc atc acc tcc aag gct gag ctg gca gag gtg agg Ala Gly Ala Ile Ser Ile Thr Ser Lys Ala Glu Leu Ala Glu Val Arg 480 485 490 495	1548		
cga gaa tgg gcc aag tac atg gaa gtc cat gag aag gcc tcc ttc acc Arg Glu Trp Ala Lys Tyr Met Glu Val His Glu Lys Ala Ser Phe Thr 500 505 510	1596		
aac agt gag ctg cac cgt gcc atg aac ctg cac gtc ggc aac ctg cgc Asn Ser Glu Leu His Arg Ala Met Asn Leu His Val Gly Asn Leu Arg 515 520 525	1644		
ctg ctc agc ggg ccg ctt gac cag gtc cgg gct gcc ctg ccc aca ccg Leu Leu Ser Gly Pro Leu Asp Gln Val Arg Ala Ala Leu Pro Thr Pro 530 535 540	1692		

gcc ctc tcc cca gag gac aag gcc gtg ctg caa aac cta aag cgc atc 1740
 Ala Leu Ser Pro Glu Asp Lys Ala Val Leu Gln Asn Leu Lys Arg Ile
 545 550 555

ctg gct aag gtg cag gag atg cgg gac cag cgc gtg tcc ctg gag cag 1788
 Leu Ala Lys Val Gln Glu Met Arg Asp Gln Arg Val Ser Leu Glu Gln
 560 565 570 575

cag ctg cgt gag ctt atc cag aaa gat gac atc act gcc tcg ctg gtc 1836
 Gln Leu Arg Glu Leu Ile Gln Lys Asp Asp Ile Thr Ala Ser Leu Val
 580 585 590

acc aca gac cac tca gag atg aag aag ttg ttc gag gag cag ctg aaa 1884
 Thr Thr Asp His Ser Glu Met Lys Lys Leu Phe Glu Glu Gln Leu Lys
 595 600 605

aag tat gac cag ctg aag gtg tac ctg gag cag aac ctg gcc gcc cag 1932
 Lys Tyr Asp Gln Leu Lys Val Tyr Leu Glu Gln Asn Leu Ala Ala Gln
 610 615 620

gac cgt gtc ctc tgt gca ctg aca gag gcc aac gtg cag tac gca gcc 1980
 Asp Arg Val Leu Cys Ala Leu Thr Glu Ala Asn Val Gln Tyr Ala Ala
 625 630 635

gtg cgg cgg gta ctc agc gac ttg gac caa aag tgg aac tcc acg ctg 2028
 Val Arg Arg Val Leu Ser Asp Leu Asp Gln Lys Trp Asn Ser Thr Leu
 640 645 650 655

cag acc ctg gtg gcc tcg tat gaa gcc tat gag gac ctg atg aag aag 2076
 Gln Thr Leu Val Ala Ser Tyr Glu Ala Tyr Glu Asp Leu Met Lys Lys
 660 665 670

tcg cag gag ggc agg gac ttc tac gca gat ctg gag agc aag gtg gct 2124
 Ser Gln Glu Gly Arg Asp Phe Tyr Ala Asp Leu Glu Ser Lys Val Ala
 675 680 685

gct ctg ctg gag cgc acg cag tcc acc tgc cag gcc cgc gag gct gcc 2172
 Ala Leu Leu Glu Arg Thr Gln Ser Thr Cys Gln Ala Arg Glu Ala Ala
 690 695 700

cgc cag cag ctc ctg gac agg gag ctg aag aag aag ccg ccg cca cgg 2220
 Arg Gln Gln Leu Leu Asp Arg Glu Leu Lys Lys Lys Pro Pro Pro Arg

705	710	715	
ccc aca gcc cca aag ccg ctg ctg ccc cgc agg gag gag agt gag gca			2268
Pro Thr Ala Pro Lys Pro Leu Leu Pro Arg Arg Glu Glu Ser Glu Ala			
720	725	730	735
gtg gaa gca gga gac ccc cct gag gag ctg cgc agc ctc ccc cct gac			2316
Val Glu Ala Gly Asp Pro Pro Glu Glu Leu Arg Ser Leu Pro Pro Asp			
740	745		750
atg gtg gct ggc cca cga ctg cct gac acc ttc ctg gga agt gcc acc			2364
Met Val Ala Gly Pro Arg Leu Pro Asp Thr Phe Leu Gly Ser Ala Thr			
755	760		765
ccg ctc cac ttt cct ccc agc ccc ttc ccc agc tcc aca ggc cca gga			2412
Pro Leu His Phe Pro Pro Ser Pro Phe Pro Ser Ser Thr Gly Pro Gly			
770	775		780
ccc cac tat ctc tca ggc ccc ttg ccc cct ggt acc tac tcg ggc ccc			2460
Pro His Tyr Leu Ser Gly Pro Leu Pro Pro Gly Thr Tyr Ser Gly Pro			
785	790		795
acc cag ctg ata cag ccc agg gcc cca ggg ccc cat gca atg ccc gta			2508
Thr Gln Leu Ile Gln Pro Arg Ala Pro Gly Pro His Ala Met Pro Val			
800	805	810	815
gca cct ggg cct gcc ctc tac cca gcc cct gca tac aca ccg gag ctg			2556
Ala Pro Gly Pro Ala Leu Tyr Pro Ala Pro Ala Tyr Thr Pro Glu Leu			
820	825		830
ggc ctt gtg ccc cga tcc tcc cca cag cat ggc gtg gtg agc agt ccc			2604
Gly Leu Val Pro Arg Ser Ser Pro Gln His Gly Val Val Ser Ser Pro			
835	840		845
tat gtg ggg gta ggg ccg gcc cca cca gtt gca ggt ctc ccc tcg gcc			2652
Tyr Val Gly Val Gly Pro Ala Pro Pro Val Ala Gly Leu Pro Ser Ala			
850	855		860
cca cct cct caa ttc tca ggc ccc gag ttg gcc atg gcg gtt cgg cca			2700
Pro Pro Pro Gln Phe Ser Gly Pro Glu Leu Ala Met Ala Val Arg Pro			
865	870		875

gcc acc acc aca gta gat agc atc cag gcg ccc atc ccc agc cac aca 2748
 Ala Thr Thr Thr Val Asp Ser Ile Gln Ala Pro Ile Pro Ser His Thr
 880 885 890 895

gcc cca cgg cca aac ccc acc cct gct cct ccc ccg ccc tgc ttc cct 2796
 Ala Pro Arg Pro Asn Pro Thr Pro Ala Pro Pro Pro Pro Cys Phe Pro
 900 905 910

gtg ccc cca ccg cag cca ctg ccc acg cct tac acc tac cct gca ggg 2844
 Val Pro Pro Pro Gln Pro Leu Pro Thr Pro Tyr Thr Tyr Pro Ala Gly
 915 920 925

gct aag caa ccc atc cca gca cag cac cac ttc tct tct ggg atc ccc 2892
 Ala Lys Gln Pro Ile Pro Ala Gln His His Phe Ser Ser Gly Ile Pro
 930 935 940

gca ggt ttt cca gcc cca agg att ggg ccc cag ccc cag ccc cat cct 2940
 Ala Gly Phe Pro Ala Pro Arg Ile Gly Pro Gln Pro Gln Pro His Pro
 945 950 955

cag ccc cat cct tca caa gcg ttt ggg cct cag ccc cca cag cag ccc 2988
 Gln Pro His Pro Ser Gln Ala Phe Gly Pro Gln Pro Pro Gln Gln Pro
 960 965 970 975

ctt cca ctc cag cat cca cat ctc ttc cca ccc cag gcc cca gga ctc 3036
 Leu Pro Leu Gln His Pro His Leu Phe Pro Pro Gln Ala Pro Gly Leu
 980 985 990

cta ccc cca caa tcc ccc tac ccc tat gcc cct cag cct ggg gtc ctg 3084
 Leu Pro Pro Gln Ser Pro Tyr Pro Tyr Ala Pro Gln Pro Gly Val Leu
 995 1000 1005

ggg cag ccg cca ccc ccc cta cac acc cag ctc tac cca ggt ccc gct 3132
 Gly Gln Pro Pro Pro Pro Leu His Thr Gln Leu Tyr Pro Gly Pro Ala
 1010 1015 1020

caa gac cct ctg cca gcc cac tca ggg gct ctg cct ttc ccc agc cct 3180
 Gln Asp Pro Leu Pro Ala His Ser Gly Ala Leu Pro Phe Pro Ser Pro
 1025 1030 1035

ggg ccc cct cag cct ccc cat ccc cca ctg gca tat ggt cct gcc cct 3228
 Gly Pro Pro Gln Pro Pro His Pro Pro Leu Ala Tyr Gly Pro Ala Pro

1040	1045	1050	1055	
tct acc aga ccc atg ggc ccc cag gca gcc cct ctt acc att cga ggg				3276
Ser Thr Arg Pro Met Gly Pro Gln Ala Ala Pro Leu Thr Ile Arg Gly				
1060	1065	1070		
ccc tcg tct gct ggc cag tcc acc cct agt ccc cac ctg gtg cct tca				3324
Pro Ser Ser Ala Gly Gln Ser Thr Pro Ser Pro His Leu Val Pro Ser				
1075	1080	1085		
cct gcc cca tct cca ggg cct ggt ccg gta ccc cct cgc ccc cca gca				3372
Pro Ala Pro Ser Pro Gly Pro Gly Pro Val Pro Pro Arg Pro Pro Ala				
1090	1095	1100		
gca gaa cca ccc cct tgc ctg cgc cga ggc gcc gca gct gca gac ctg				3420
Ala Glu Pro Pro Pro Cys Leu Arg Arg Gly Ala Ala Ala Asp Leu				
1105	1110	1115		
ctc tcc tcc agc ccg gag agc cag cat ggc ggc act cag tct cct ggg				3468
Leu Ser Ser Ser Pro Glu Ser Gln His Gly Gly Thr Gln Ser Pro Gly				
1120	1125	1130	1135	
ggt ggg cag ccc ctg ctg cag ccc acc aag gtg gat gca gct gag ggt				3516
Gly Gly Gln Pro Leu Leu Gln Pro Thr Lys Val Asp Ala Ala Glu Gly				
1140	1145	1150		
cgt cgg ccg cag gcc ctg cgg ctg att gag cgg gac ccc tat gag cat				3564
Arg Arg Pro Gln Ala Leu Arg Leu Ile Glu Arg Asp Pro Tyr Glu His				
1155	1160	1165		
cct gag agg ctg cgg cag ttg cag cag gag ctg gag gcc ttt cgg ggt				3612
Pro Glu Arg Leu Arg Gln Leu Gln Gln Glu Leu Glu Ala Phe Arg Gly				
1170	1175	1180		
cag ctg ggg gat gtg gga gct ctg gac act gtc tgg cga gag ctg caa				3660
Gln Leu Gly Asp Val Gly Ala Leu Asp Thr Val Trp Arg Glu Leu Gln				
1185	1190	1195		
gat gcg cag gaa cat gat gcc cga ggc cgt tcc atc gcc att gcc cgc				3708
Asp Ala Gln Glu His Asp Ala Arg Gly Arg Ser Ile Ala Ile Ala Arg				
1200	1205	1210	1215	

tgc tac tca ctg aag aac cgg cac cag gat gtc atg ccc tat gac agt 3756
 Cys Tyr Ser Leu Lys Asn Arg His Gln Asp Val Met Pro Tyr Asp Ser
 1220 1225 1230

aac cgt gtg gtg ctg cgc tca ggc aag gat gac tac atc aat gcc agc 3804
 Asn Arg Val Val Leu Arg Ser Gly Lys Asp Asp Tyr Ile Asn Ala Ser
 1235 1240 1245

tgc gtg gag ggg ctc tcc cca tac tgc ccc ccg cta gtg gca acc cag 3852
 Cys Val Glu Gly Leu Ser Pro Tyr Cys Pro Pro Leu Val Ala Thr Gln
 1250 1255 1260

gcc cca ctg cct ggc aca gct gct gac ttc tgg ctc atg gtc cat gag 3900
 Ala Pro Leu Pro Gly Thr Ala Ala Asp Phe Trp Leu Met Val His Glu
 1265 1270 1275

cag aaa gtg tca gtc att gtc atg ctg gtt tct gag gct gag atg gag 3948
 Gln Lys Val Ser Val Ile Val Met Leu Val Ser Glu Ala Glu Met Glu
 1280 1285 1290 1295

aag caa aaa gtg gca cgc tac ttc ccc acc gag agg ggc cag ccc atg 3996
 Lys Gln Lys Val Ala Arg Tyr Phe Pro Thr Glu Arg Gly Gln Pro Met
 1300 1305 1310

gtg cac ggt gcc ctg agc ctg gca ttg agc agc gtc cgc agc acc gaa 4044
 Val His Gly Ala Leu Ser Leu Ala Leu Ser Ser Val Arg Ser Thr Glu
 1315 1320 1325

acc cat gtg gag cgc gtg ctg agc ctg cag ttc cga gac cag agc ctc 4092
 Thr His Val Glu Arg Val Leu Ser Leu Gln Phe Arg Asp Gln Ser Leu
 1330 1335 1340

aag cgc tct ctt gtg cac ctg cac ttc ccc act tgg cct gag tta ggc 4140
 Lys Arg Ser Leu Val His Leu His Phe Pro Thr Trp Pro Glu Leu Gly
 1345 1350 1355

ctg ccc gac agc ccc agc aac ttg ctg cgc ttc atc cag gag gtg cac 4188
 Leu Pro Asp Ser Pro Ser Asn Leu Leu Arg Phe Ile Gln Glu Val His
 1360 1365 1370 1375

gca cat tac ctg cat cag cgg ccg ctg cac acg ccc atc att gtg cac 4236
 Ala His Tyr Leu His Gln Arg Pro Leu His Thr Pro Ile Ile Val His

1380	1385	1390	
tgc agc tct ggt gtg ggc cgc acg gga gcc ttt gca ctg ctc tat gca			4284
Cys Ser Ser Gly Val Gly Arg Thr Gly Ala Phe Ala Leu Leu Tyr Ala			
1395	1400	1405	
gct gtg cag gag gtg gag gct ggg aac gga atc cct gag ctg cct cag			4332
Ala Val Gln Glu Val Glu Ala Gly Asn Gly Ile Pro Glu Leu Pro Gln			
1410	1415	1420	
ctg gtg cgg cgc atg cgg cag cag aga aag cac atg ctg cag gag aag			4380
Leu Val Arg Arg Met Arg Gln Gln Arg Lys His Met Leu Gln Glu Lys			
1425	1430	1435	
ctg cac ctc agg ttc tgc tat gag gca gtg gtg aga cac gtg gag cag			4428
Leu His Leu Arg Phe Cys Tyr Glu Ala Val Val Arg His Val Glu Gln			
1440	1445	1450	1455
gtc ctg cag cgc cat ggt gtg cct cct cca tgc aaa ccc ttg gcc agt			4476
Val Leu Gln Arg His Gly Val Pro Pro Pro Cys Lys Pro Leu Ala Ser			
1460	1465	1470	
gca agc atc agc cag aag aac cac ctt cct cag gac tcc cag gac ctg			4524
Ala Ser Ile Ser Gln Lys Asn His Leu Pro Gln Asp Ser Gln Asp Leu			
1475	1480	1485	
gtc ctc ggt ggg gat gtg ccc atc agc tcc atc cag gcc acc att gcc			4572
Val Leu Gly Gly Asp Val Pro Ile Ser Ser Ile Gln Ala Thr Ile Ala			
1490	1495	1500	
aag ctc agc att cgg cct cct ggg ggg ttg gag tcc ccg gtt gcc agc			4620
Lys Leu Ser Ile Arg Pro Pro Gly Gly Leu Glu Ser Pro Val Ala Ser			
1505	1510	1515	
ttg cca ggc cct gca gag ccc cca ggc ctc ccg cca gcc agc ctc cca			4668
Leu Pro Gly Pro Ala Glu Pro Pro Gly Leu Pro Pro Ala Ser Leu Pro			
1520	1525	1530	1535
gag tct acc cca atc cca tct tcc tcc cca ccc ccc ctt tcc tcc cca			4716
Glu Ser Thr Pro Ile Pro Ser Ser Ser Pro Pro Pro Leu Ser Ser Pro			
1540	1545	1550	

cta cct gag gct ccc cag cct aag gag gag ccg cca gtg cct gaa gcc 4764
 Leu Pro Glu Ala Pro Gln Pro Lys Glu Glu Pro Pro Val Pro Glu Ala
 1555 1560 1565

ccc agc tcg ggg ccc ccc tcc tcc tcc ctg gaa ttg ctg gcc tcc ttg 4812
 Pro Ser Ser Gly Pro Pro Ser Ser Ser Leu Glu Leu Leu Ala Ser Leu
 1570 1575 1580

acc cca gag gcc ttc tcc ctg gac agc tcc ctg cgg ggc aaa cag cgg 4860
 Thr Pro Glu Ala Phe Ser Leu Asp Ser Ser Leu Arg Gly Lys Gln Arg
 1585 1590 1595

atg agc aag cat aac ttt ctg cag gcc cat aac ggg caa ggg ctg cgg 4908
 Met Ser Lys His Asn Phe Leu Gln Ala His Asn Gly Gln Gly Leu Arg
 1600 1605 1610 1615

gcc acc cgg ccc tct gac gac ccc ctc agc ctt ctg gat cca ctc tgg 4956
 Ala Thr Arg Pro Ser Asp Asp Pro Leu Ser Leu Leu Asp Pro Leu Trp
 1620 1625 1630

aca ctc aac aag acc tgaacagggtt ttgcctacct ggctcctaca ctacatcatc 5011
 Thr Leu Asn Lys Thr
 1635

atcatctcat gccacactgc ccaacaccag cagagcttct cagtgggcac agtctcttac 5071

tcccatttct gctgcctttg gccctgectg gccagcctg caccctgtg ggggtggaaat 5131

gtactgcagg ctctgggtca gggtctgctc ctttatggga cccgacattt ttcagctctt 5191

tgctattgaa ataataaacc accctgttct gtgaaaaaaaa aaa 5234

<210> 2

<211> 1636

<212> PRT

<213> Homo sapiens

<400> 2

Met Glu Ala Val Pro Arg Met Pro Val Ile Trp Leu Asp Leu Lys Glu
 1 5 10 15

Ala Gly Asp Phe His Phe Gln Pro Ala Val Lys Lys Phe Val Leu Lys
 20 25 30

Asn Tyr Gly Glu Asn Pro Glu Ala Tyr Asn Glu Glu Leu Lys Lys Leu
 35 40 45

Glu Leu Leu Arg Gln Asn Ala Val Arg Val Pro Arg Asp Phe Glu Gly
 50 55 60

Cys Ser Val Leu Arg Lys Tyr Leu Gly Gln Leu His Tyr Leu Gln Ser
 65 70 75 80

Arg Val Pro Met Gly Ser Gly Gln Glu Ala Ala Val Pro Val Thr Trp
 85 90 95

Thr Glu Ile Phe Ser Gly Lys Ser Val Ala His Glu Asp Ile Lys Tyr
 100 105 110

Glu Gln Ala Cys Ile Leu Tyr Asn Leu Gly Ala Leu His Ser Met Leu
 115 120 125

Gly Ala Met Asp Lys Arg Val Ser Glu Glu Gly Met Lys Val Ser Cys
 130 135 140

Thr His Phe Gln Cys Ala Ala Gly Ala Phe Ala Tyr Leu Arg Glu His
 145 150 155 160

Phe Pro Gln Ala Tyr Ser Val Asp Met Ser Arg Gln Ile Leu Thr Leu
 165 170 175

Asn Val Asn Leu Met Leu Gly Gln Ala Gln Glu Cys Leu Leu Glu Lys
 180 185 190

Ser Met Leu Asp Asn Arg Lys Ser Phe Leu Val Ala Arg Ile Ser Ala
 195 200 205

Gln Val Val Asp Tyr Tyr Lys Glu Ala Cys Arg Ala Leu Glu Asn Pro
 210 215 220

Asp Thr Ala Ser Leu Leu Gly Arg Ile Gln Lys Asp Trp Lys Lys Leu
 225 230 235 240

Val Gln Met Lys Ile Tyr Tyr Phe Ala Ala Val Ala His Leu His Met
 245 250 255
 Gly Lys Gln Ala Glu Glu Gln Gln Lys Phe Gly Glu Arg Val Ala Tyr
 260 265 270
 Phe Gln Ser Ala Leu Asp Lys Pro Asn Glu Ala Ile Lys Leu Ala Lys
 275 280 285
 Gly Gln Pro Asp Thr Val Gln Asp Ala Leu Arg Phe Thr Met Asp Val
 290 295 300
 Ile Gly Gly Lys Tyr Asn Ser Ala Lys Lys Asp Asn Asp Phe Ile Tyr
 305 310 315 320
 His Glu Ala Val Pro Ala Leu Asp Thr Leu Gln Pro Val Lys Gly Ala
 325 330 335
 Pro Leu Val Lys Pro Leu Pro Val Asn Pro Thr Asp Pro Ala Val Thr
 340 345 350
 Gly Pro Asp Ile Phe Ala Lys Leu Val Pro Met Ala Ala His Glu Ala
 355 360 365
 Ser Ser Leu Tyr Ser Glu Glu Lys Ala Lys Leu Leu Arg Glu Met Met
 370 375 380
 Ala Lys Ile Glu Asp Lys Asn Glu Val Leu Asp Gln Phe Met Asp Ser
 385 390 395 400
 Met Gln Leu Asp Pro Glu Thr Val Asp Asn Leu Asp Ala Tyr Ser His
 405 410 415
 Ile Pro Pro Gln Leu Met Glu Lys Cys Ala Ala Leu Ser Val Arg Pro
 420 425 430
 Asp Thr Val Arg Asn Leu Val Gln Ser Met Gln Val Leu Ser Gly Val
 435 440 445
 Phe Thr Asp Val Glu Ala Ser Leu Lys Asp Ile Arg Asp Leu Leu Glu
 450 455 460

Glu Asp Glu Leu Leu Glu Gln Lys Phe Gln Glu Ala Val Gly Gln Ala
 465 470 475 480

Gly Ala Ile Ser Ile Thr Ser Lys Ala Glu Leu Ala Glu Val Arg Arg
 485 490 495

Glu Trp Ala Lys Tyr Met Glu Val His Glu Lys Ala Ser Phe Thr Asn
 500 505 510

Ser Glu Leu His Arg Ala Met Asn Leu His Val Gly Asn Leu Arg Leu
 515 520 525

Leu Ser Gly Pro Leu Asp Gln Val Arg Ala Ala Leu Pro Thr Pro Ala
 530 535 540

Leu Ser Pro Glu Asp Lys Ala Val Leu Gln Asn Leu Lys Arg Ile Leu
 545 550 555 560

Ala Lys Val Gln Glu Met Arg Asp Gln Arg Val Ser Leu Glu Gln Gln
 565 570 575

Leu Arg Glu Leu Ile Gln Lys Asp Asp Ile Thr Ala Ser Leu Val Thr
 580 585 590

Thr Asp His Ser Glu Met Lys Lys Leu Phe Glu Glu Gln Leu Lys Lys
 595 600 605

Tyr Asp Gln Leu Lys Val Tyr Leu Glu Gln Asn Leu Ala Ala Gln Asp
 610 615 620

Arg Val Leu Cys Ala Leu Thr Glu Ala Asn Val Gln Tyr Ala Ala Val
 625 630 635 640

Arg Arg Val Leu Ser Asp Leu Asp Gln Lys Trp Asn Ser Thr Leu Gln
 645 650 655

Thr Leu Val Ala Ser Tyr Glu Ala Tyr Glu Asp Leu Met Lys Lys Ser
 660 665 670

Gln Glu Gly Arg Asp Phe Tyr Ala Asp Leu Glu Ser Lys Val Ala Ala
 675 680 685

Leu Leu Glu Arg Thr Gln Ser Thr Cys Gln Ala Arg Glu Ala Ala Arg
 690 695 700

Gln Gln Leu Leu Asp Arg Glu Leu Lys Lys Lys Pro Pro Pro Arg Pro
 705 710 715 720

Thr Ala Pro Lys Pro Leu Leu Pro Arg Arg Glu Glu Ser Glu Ala Val
 725 730 735

Glu Ala Gly Asp Pro Pro Glu Glu Leu Arg Ser Leu Pro Pro Asp Met
 740 745 750

Val Ala Gly Pro Arg Leu Pro Asp Thr Phe Leu Gly Ser Ala Thr Pro
 755 760 765

Leu His Phe Pro Pro Ser Pro Phe Pro Ser Ser Thr Gly Pro Gly Pro
 770 775 780

His Tyr Leu Ser Gly Pro Leu Pro Pro Gly Thr Tyr Ser Gly Pro Thr
 785 790 795 800

Gln Leu Ile Gln Pro Arg Ala Pro Gly Pro His Ala Met Pro Val Ala
 805 810 815

Pro Gly Pro Ala Leu Tyr Pro Ala Pro Ala Tyr Thr Pro Glu Leu Gly
 820 825 830

Leu Val Pro Arg Ser Ser Pro Gln His Gly Val Val Ser Ser Pro Tyr
 835 840 845

Val Gly Val Gly Pro Ala Pro Pro Val Ala Gly Leu Pro Ser Ala Pro
 850 855 860

Pro Pro Gln Phe Ser Gly Pro Glu Leu Ala Met Ala Val Arg Pro Ala
 865 870 875 880

Thr Thr Thr Val Asp Ser Ile Gln Ala Pro Ile Pro Ser His Thr Ala
 885 890 895

Pro Arg Pro Asn Pro Thr Pro Ala Pro Pro Pro Pro Cys Phe Pro Val
 900 905 910

Pro Pro Pro Gln Pro Leu Pro Thr Pro Tyr Thr Tyr Pro Ala Gly Ala
 915 920 925

Lys Gln Pro Ile Pro Ala Gln His His Phe Ser Ser Gly Ile Pro Ala
 930 935 940

Gly Phe Pro Ala Pro Arg Ile Gly Pro Gln Pro Gln Pro His Pro Gln
 945 950 955 960

Pro His Pro Ser Gln Ala Phe Gly Pro Gln Pro Pro Gln Gln Pro Leu
 965 970 975

Pro Leu Gln His Pro His Leu Phe Pro Pro Gln Ala Pro Gly Leu Leu
 980 985 990

Pro Pro Gln Ser Pro Tyr Pro Tyr Ala Pro Gln Pro Gly Val Leu Gly
 995 1000 1005

Gln Pro Pro Pro Pro Leu His Thr Gln Leu Tyr Pro Gly Pro Ala Gln
 1010 1015 1020

Asp Pro Leu Pro Ala His Ser Gly Ala Leu Pro Phe Pro Ser Pro Gly
 1025 1030 1035 104

Pro Pro Gln Pro Pro His Pro Pro Leu Ala Tyr Gly Pro Ala Pro Ser
 1045 1050 1055

Thr Arg Pro Met Gly Pro Gln Ala Ala Pro Leu Thr Ile Arg Gly Pro
 1060 1065 1070

Ser Ser Ala Gly Gln Ser Thr Pro Ser Pro His Leu Val Pro Ser Pro
 1075 1080 1085

Ala Pro Ser Pro Gly Pro Gly Pro Val Pro Pro Arg Pro Pro Ala Ala
 1090 1095 1100

Glu Pro Pro Pro Cys Leu Arg Arg Gly Ala Ala Ala Ala Asp Leu Leu
 1105 1110 1115 112

Ser Ser Ser Pro Glu Ser Gln His Gly Gly Thr Gln Ser Pro Gly Gly
 1125 1130 1135

Gly Gln Pro Leu Leu Gln Pro Thr Lys Val Asp Ala Ala Glu Gly Arg
 1140 1145 1150
 Arg Pro Gln Ala Leu Arg Leu Ile Glu Arg Asp Pro Tyr Glu His Pro
 1155 1160 1165
 Glu Arg Leu Arg Gln Leu Gln Gln Glu Leu Glu Ala Phe Arg Gly Gln
 1170 1175 1180
 Leu Gly Asp Val Gly Ala Leu Asp Thr Val Trp Arg Glu Leu Gln Asp
 185 1190 1195 120
 Ala Gln Glu His Asp Ala Arg Gly Arg Ser Ile Ala Ile Ala Arg Cys
 1205 1210 1215
 Tyr Ser Leu Lys Asn Arg His Gln Asp Val Met Pro Tyr Asp Ser Asn
 1220 1225 1230
 Arg Val Val Leu Arg Ser Gly Lys Asp Asp Tyr Ile Asn Ala Ser Cys
 1235 1240 1245
 Val Glu Gly Leu Ser Pro Tyr Cys Pro Pro Leu Val Ala Thr Gln Ala
 1250 1255 1260
 Pro Leu Pro Gly Thr Ala Ala Asp Phe Trp Leu Met Val His Glu Gln
 265 1270 1275 128
 Lys Val Ser Val Ile Val Met Leu Val Ser Glu Ala Glu Met Glu Lys
 1285 1290 1295
 Gln Lys Val Ala Arg Tyr Phe Pro Thr Glu Arg Gly Gln Pro Met Val
 1300 1305 1310
 His Gly Ala Leu Ser Leu Ala Leu Ser Ser Val Arg Ser Thr Glu Thr
 1315 1320 1325
 His Val Glu Arg Val Leu Ser Leu Gln Phe Arg Asp Gln Ser Leu Lys
 1330 1335 1340
 Arg Ser Leu Val His Leu His Phe Pro Thr Trp Pro Glu Leu Gly Leu
 345 1350 1355 136

Pro Asp Ser Pro Ser Asn Leu Leu Arg Phe Ile Gln Glu Val His Ala
 1365 1370 1375
 His Tyr Leu His Gln Arg Pro Leu His Thr Pro Ile Ile Val His Cys
 1380 1385 1390
 Ser Ser Gly Val Gly Arg Thr Gly Ala Phe Ala Leu Leu Tyr Ala Ala
 1395 1400 1405
 Val Gln Glu Val Glu Ala Gly Asn Gly Ile Pro Glu Leu Pro Gln Leu
 1410 1415 1420
 Val Arg Arg Met Arg Gln Gln Arg Lys His Met Leu Gln Glu Lys Leu
 1425 1430 1435 1444
 His Leu Arg Phe Cys Tyr Glu Ala Val Val Arg His Val Glu Gln Val
 1445 1450 1455
 Leu Gln Arg His Gly Val Pro Pro Pro Cys Lys Pro Leu Ala Ser Ala
 1460 1465 1470
 Ser Ile Ser Gln Lys Asn His Leu Pro Gln Asp Ser Gln Asp Leu Val
 1475 1480 1485
 Leu Gly Gly Asp Val Pro Ile Ser Ser Ile Gln Ala Thr Ile Ala Lys
 1490 1495 1500
 Leu Ser Ile Arg Pro Pro Gly Gly Leu Glu Ser Pro Val Ala Ser Leu
 505 1510 1515 152
 Pro Gly Pro Ala Glu Pro Pro Gly Leu Pro Pro Ala Ser Leu Pro Glu
 1525 1530 1535
 Ser Thr Pro Ile Pro Ser Ser Ser Pro Pro Pro Leu Ser Ser Pro Leu
 1540 1545 1550
 Pro Glu Ala Pro Gln Pro Lys Glu Glu Pro Pro Val Pro Glu Ala Pro
 1555 1560 1565
 Ser Ser Gly Pro Pro Ser Ser Ser Leu Glu Leu Leu Ala Ser Leu Thr
 1570 1575 1580

Pro Glu Ala Phe Ser Leu Asp Ser Ser Leu Arg Gly Lys Gln Arg Met
585 1590 1595 160

Ser Lys His Asn Phe Leu Gln Ala His Asn Gly Gln Gly Leu Arg Ala
1605 1610 1615

Thr Arg Pro Ser Asp Asp Pro Leu Ser Leu Leu Asp Pro Leu Trp Thr
1620 1625 1630

Leu Asn Lys Thr
1635

<210> 3

<211> 9309

<212> DNA

<213> Homo sapiens

<220>

<221> exon

<222> (529)..(656)

<220>

<221> exon

<222> (881)..(959)

<220>

<221> exon

<222> (1625)..(1674)

<220>

<221> exon

<222> (1791)..(1922)

<220>

<221> exon

<222> (2201)..(2281)

<220>

<221> exon

<222> (2357)..(2488)

<220>

<221> exon

<222> (2579)..(2626)

<220>

<221> exon

<222> (3006)..(3062)

<220>

<221> exon

<222> (3185)..(3243)

<220>

<221> exon

<222> (3381)..(3460)

<220>

<221> exon

<222> (3573)..(3687)

<220>

<221> exon

<222> (3766)..(3831)

<220>

<221> exon

<222> (4221)..(4366)

<220>

<221> exon

<222> (4652)..(4963)

<220>

<221> exon
<222> (5039)..(5193)

<220>
<221> exon
<222> (5293)..(5444)

<220>
<221> exon
<222> (5531)..(5710)

<220>
<221> exon
<222> (5804)..(7562)

<220>
<221> exon
<222> (7657)..(7841)

<220>
<221> exon
<222> (7968)..(8072)

<220>
<221> exon
<222> (8157)..(8295)

<220>
<221> exon
<222> (8388)..(8501)

<220>
<221> exon
<222> (8580)..(9309)

<220>
<221> intron

<222> (1)..(528)

<220>

<221> intron

<222> (657)..(880)

<220>

<221> intron

<222> (958)..(1624)

<220>

<221> intron

<222> (1675)..(1790)

<220>

<221> intron

<222> (1923)..(2200)

<220>

<221> intron

<222> (2282)..(2356)

<220>

<221> intron

<222> (2488)..(2578)

<220>

<221> intron

<222> (2627)..(3005)

<220>

<221> intron

<222> (3063)..(3184)

<220>

<221> intron

<222> (3244)..(3380)

<220>

<221> intron

<222> (3461)..(3572)

<220>

<221> intron

<222> (3688)..(3765)

<220>

<221> intron

<222> (3832)..(4220)

<220>

<221> intron

<222> (4367)..(4651)

<220>

<221> intron

<222> (4964)..(5038)

<220>

<221> intron

<222> (5194)..(5292)

<220>

<221> intron

<222> (5445)..(5530)

<220>

<221> intron

<222> (5711)..(5803)

<220>

<221> intron

<222> (7563)..(7656)

<220>

<221> intron

<222> (7842)..(7967)

<220>

<221> intron

<222> (8073)..(8156)

<220>

<221> intron

<222> (8296)..(8387)

<220>

<221> intron

<222> (8502)..(8579)

<400> 42

gttgcaaatg ttttcagatg tcttgagatg gggctgggca cagtgttca tgcctgtaat 60
cccagcactt cgagaggctg aggcgggtgg atcacaagtt caagtgttcg agaccagcct 120
ggccaaaatg ggaaaccgc cttctactaa aaatacaaaa ttagccaggc atggtggcgt 180
gtgcctgtag tctcagctaa ttgggaggct gaggcaggag aattgcttga acccaggagg 240
ttgcagtga tctgcattc cagcctgggt gacagaatga gactcttttt ttttttttt 300
tttaaaaaag gaaaaactat gtgggataga ttcagtcggc cagctgccag cctggcttgt 360
ggaatgtttt gttgaacgac tataggtata gccatggact ttttgggatc ccttggtgtg 420
tgcagtcggc cagctgtgat gcattttcct ggcatttgtg agttagctgg cctgcgtgtc 480
cactgcccc tatgacaaca ggcctgcttc tccatcctg cccccag aat gct gtc 537
Asn Ala Val

1

cgt gtc cca cga gac ttt gag ggc tgt agt gtc ctc cgc aag tac ctc 585
 Arg Val Pro Arg Asp Phe Glu Gly Cys Ser Val Leu Arg Lys Tyr Leu
 5 10 15

ggc cag ctt cat tac ctg cag agt cgg gtc ccc atg ggc tcg ggc cag 633
 Gly Gln Leu His Tyr Leu Gln Ser Arg Val Pro Met Gly Ser Gly Gln
 20 25 30 35

gag gcc gct gtc cct gtc acc tg gtgagagccg caggcagggc tggaggatcc 686
 Glu Ala Ala Val Pro Val Thr Trp
 40

cacggggagt ctgggtggtg ggggtgctcc tccccctcct tcccccttct ccttgctgca 746

taggatggga atggcctcat atggtcccc caggcctccc cacatcccca caatgggggtg 806

tttagtcccc ttacctaate tcagggccct ggctagctcc tgcccttgag ataagtgggg 866

tgccatgtct gcag g aca gag atc ttc tca ggc aag tct gtg gcc cat gag 917
 Thr Glu Ile Phe Ser Gly Lys Ser Val Ala His Glu
 45 50 55

gac atc aag tac gag cag gcc tgt att ctc tac aac ctt g gtgagctgcc 967
 Asp Ile Lys Tyr Glu Gln Ala Cys Ile Leu Tyr Asn Leu
 60 65

tgatcccttc ccccgccct actccccagt cctgccagcc tagctttcag ctcttcagat 1027

ggccacaaag gcagtgggct gataaaggga ggactccagg attcccgccc ctccactgac 1087

ctccccacag ccttgccagc tcctccactg ttttctgggc tgggcccgtg gggagcctct 1147

gctggggcga ggctctactg ggctggctgc cacacagagt tgccactgtg tgctctgtct 1207

gggagagagg gcctttcttc tttgactgga cagccctcc ccaactgtgtt tttgggctgc 1267

ttcaggcagg aggagagtgt ggctggcctc agcttacct gctagtcagc cttagcctgg 1327

ctcaagggag aagcttgggg agccccagag ttctgtgcaa acatccctga agcttcaaga 1387
 ttggaccctt tggacaggag cccttccatc ctctgccaaa tgcagagcag gagactgagg 1447
 gctgggcagg acttctgagt ttccctgttc cctccagctt ggggtgtcctt ggactcgttg 1507
 ccctggttct cagagccatg ttgtctgat tgtggataac agcagtgcc cctctgtctc 1567
 accttcacat ggggtgtgagc agccccaggc ccctaact gtcctctccc tccccag ga 1626
 Gly
 gcg ctg cac tcc atg ctg ggg gcc atg gac aag cgg gtg tct gag gag 1674
 Ala Leu His Ser Met Leu Gly Ala Met Asp Lys Arg Val Ser Glu Glu
 70 75 80 85
 gtgaggagag gggcagtagt ggaacatgtg gacataccag ggaggggcag cctcccaagt 1734
 atggatgaat cctgacccat ggagtggaca caggccatcc tcccactccc tcccag ggc 1793
 Gly
 atg aag gtc tcc tgt acc cat ttc cag tgc gca gcc ggc gcc ttc gcc 1841
 Met Lys Val Ser Cys Thr His Phe Gln Cys Ala Ala Gly Ala Phe Ala
 90 95 100
 tac cta cgg gag cac ttc cct caa gcc tac agc gtc gac atg agc cgc 1889
 Tyr Leu Arg Glu His Phe Pro Gln Ala Tyr Ser Val Asp Met Ser Arg
 105 110 115
 cag atc ctt acg ctc aac gtc aac ctc atg ctg gtgaggagc gccctggttg 1942
 Gln Ile Leu Thr Leu Asn Val Asn Leu Met Leu
 120 125
 gagtggagtt gaatccaggg aaggggatgg ccagggaggg ggcagttggg cctggatcct 2002
 ggaccaaggc agtgaggac aagcaagggg ccttggcttt gttgaatcag gagcacggtg 2062

gtgctgcttg gagtgggggc agctggggga gagggcagtg aagagggate cctcagtcctg 2122
 cccctgggga aggataaagg gaaggggaag cgggaggcct tgggtgagcg agggagtgc 2182
 cctcacgtgt cgccccag ggc cag gct cag gag tgc ctc ctg gag aag tcg 2233
 Gly Gln Ala Gln Glu Cys Leu Leu Glu Lys Ser
 130 135 140
 atg ttg gac aac agg aag agc ttt ctg gtg gcc cgc atc agt gca cag 2281
 Met Leu Asp Asn Arg Lys Ser Phe Leu Val Ala Arg Ile Ser Ala Gln
 145 150 155
 gtaggacgg ggctgagggg aggccttcac tttactgctg actccccac tcattgggcc 2341
 ccacctgtt ctcag gtg gta gat tac tac aag gag gca tgc cgg gcc ttg 2392
 Val Val Asp Tyr Tyr Lys Glu Ala Cys Arg Ala Leu
 160 165
 gag aac ccc gac act gcc tca ctg ctg ggc cgg atc cag aag gac tgg 2440
 Glu Asn Pro Asp Thr Ala Ser Leu Leu Gly Arg Ile Gln Lys Asp Trp
 170 175 180
 aag aaa ctt gtg cag atg aag atc tac tac ttc gca gcc gtg gct cat 2488
 Lys Lys Leu Val Gln Met Lys Ile Tyr Tyr Phe Ala Ala Val Ala His
 185 190 195 200
 gtgagggcct ggggccccag ggcggggcag ggcggggctg agtggccaca gtcaggaag 2548
 caagtcgtgg cgtctcttct tctttcccag ctg cac atg gga aag cag gcc gag 2602
 Leu His Met Gly Lys Gln Ala Glu
 205
 gag cag cag aag ttc ggg gag cgg gtgagctaca gcgaggagg gactggggac 2656
 Glu Gln Gln Lys Phe Gly Glu Arg
 210 215

caatggcagc cttcagttag atgccggtgt gctcccgtc cttacacacc atggggtggc 2716
 cttctctgtc tcaccctggc catcacctg ctggaggeet ggtgtcttaa gtgtgtccc 2776
 atctgtgcag ccctcgtccc tcggggtctg aggaggtggg gcaggctcca tacagagcag 2836
 gtggctgggg cagggtgtgg cgccacctg ctgtgttgg ctggggtggt gcccggtgc 2896
 ctctgagct gcttgtcatc tgatggacag gcagggccgg gtgggaggca ggaggagaaa 2956
 aggtccaagc agaggaggac agagcaggct ttctgccac cctccacag gtt gca tac 3014
 Val Ala Tyr
 ttc cag agc gcc ctg gac aag ccc aat gaa gcc atc aag ttg gcc aag 3062
 Phe Gln Ser Ala Leu Asp Lys Pro Asn Glu Ala Ile Lys Leu Ala Lys
 220 225 230 235
 gtaaagctga ggaaggcctg gctgccctga gggtatagga gcaagcccgg taggactgag 3122
 ggggtgtcct ggtgccagcc ttggttagtg ctaaggcccc acccctgtcc ctaaccccac 3182
 ag ggc cag cct gac act gtg caa gac gcg ctt cgc ttc act atg gat 3229
 Gly Gln Pro Asp Thr Val Gln Asp Ala Leu Arg Phe Thr Met Asp
 240 245 250
 gtc att ggg gga aa gtgagtctgt gggggtggcc ctggttcct ctttttgtga 3283
 Val Ile Gly Gly Lys
 255
 aggtcttgt cctctgctg gcattacat gggaagtagg ttctggatcc ccacggacac 3343
 cccgtgactg cccactcccc ctgtcctga tccccag g tac aat tct gcc aag 3396
 Tyr Asn Ser Ala Lys
 260
 aag gac aac gac ttc att tac cat gag gct gtc cca gca ttg gac acc 3444
 Lys Asp Asn Asp Phe Ile Tyr His Glu Ala Val Pro Ala Leu Asp Thr

265	270	275	
ctt cag cct gta aaa g gtcggggagc tgagaggtgg gggcagaggt gacggtgggg 3500 Leu Gln Pro Val Lys 280			
tggggacagg acacaggagg ctgcctcaag gactctgcgt gggcctgac tccacaattc 3560 ccaccccccc ag ga gcc ccc ttg gtg aag ccc ttg cca gtg aac ccc aca 3610 Gly Ala Pro Leu Val Lys Pro Leu Pro Val Asn Pro Thr 285 290			
gac cca gct gtt aca ggc cct gac atc ttt gcc aaa ctg gta ccc atg 3658 Asp Pro Ala Val Thr Gly Pro Asp Ile Phe Ala Lys Leu Val Pro Met 295 300 305 310			
gct gcc cac gag gcc tcg tca ctg tac ag gtgggtggag ggtggcacag 3707 Ala Ala His Glu Ala Ser Ser Leu Tyr Ser 315 320			
agggaggtgg ggtgtcttga gatgtgggtc ttcagcaaat gctctgtcgc ttctgcag t 3766 gag gag aag gcc aag ctg ctc cgg gag atg atg gcc aag att gag gac 3814 Glu Glu Lys Ala Lys Leu Leu Arg Glu Met Met Ala Lys Ile Glu Asp 325 330 335			
aag aat gag gtc ctg ga gtgagtgtgg gacttgggca gggaggcgga 3861 Lys Asn Glu Val Leu Glu 340			
ggcaggcagc acttcccggg cctctggggg ccccagggt gcctatgctg ggagaggaat 3921 gaaatgtcca ttccaaacag gtttcccaat gctgccttcc cgcccggggt ggtggggcta 3981 gtgtgtaagg caggagtcat gtcttgggag gaggaggtgc cttctgttcc actgtttcca 4041 gcagtgcctt gggcatgttc tgtgagacca ggccagacct ggtagtaggg gtccgaggtc 4101			

acaatttgct ctctgctgag acctcagatt gagggtgag gcttgcccta gcggctcctt 4161
 tgacatggtc agagtggat cagcaccag caccacctg gccctgttgc tccccacag 4220
 c cag ttc atg gat tca atg cag ttg gat ccc gag acg gtg gac aac ctt 4269
 Gln Phe Met Asp Ser Met Gln Leu Asp Pro Glu Thr Val Asp Asn Leu
 345 350 355
 gat gcc tac agc cac atc cca ccc cag ctc atg gag aag tgc gcg gct 4317
 Asp Ala Tyr Ser His Ile Pro Pro Gln Leu Met Glu Lys Cys Ala Ala
 360 365 370
 ctc agc gtc cgg ccc gac act gtc agg aac ctt gta cag tcc atg caa g 4366
 Leu Ser Val Arg Pro Asp Thr Val Arg Asn Leu Val Gln Ser Met Gln
 375 380 385 390
 gtgagtaagg ggcagagcaa gcaggtgaa gggagtgtgg aggtcatcta ctgtggcctc 4426
 ctccgtgtcc ctggtcactg aggatggaga ctgcaccct ctaggccctg gcttgggcat 4486
 ccacaccac tcctctgaat cagcatacct cttgcaccct gctcagtgtg cgctgggcct 4546
 cacttaagcc ctgacctgag ggggcggttc tgtctcttgg gggaggggcc catgggtgcc 4606
 cggtcagcct gcctcagggg ctgcctgtac aatccacaac ccag tg ctg tca ggt 4662
 Val Leu Ser Gly
 gtg ttc acg gat gtg gag gct tcc ctg aag gac atc aga gat ctg ttg 4710
 Val Phe Thr Asp Val Glu Ala Ser Leu Lys Asp Ile Arg Asp Leu Leu
 395 400 405 410
 gag gag gat gag ctg cta gag cag aag ttt cag gag gcg gtg ggc cag 4758
 Glu Glu Asp Glu Leu Leu Glu Gln Lys Phe Gln Glu Ala Val Gly Gln
 415 420 425
 gca ggg gcc atc tcc atc acc tcc aag gct gag ctg gca gag gtg agg 4806

Ala Gly Ala Ile Ser Ile Thr Ser Lys Ala Glu Leu Ala Glu Val Arg
 430 435 440

cga gaa tgg gcc aag tac atg gaa gtc cat gag aag gcc tcc ttc acc 4854
 Arg Glu Trp Ala Lys Tyr Met Glu Val His Glu Lys Ala Ser Phe Thr
 445 450 455

aac agt gag ctg cac cgt gcc atg aac ctg cac gtc ggc aac ctg cgc 4902
 Asn Ser Glu Leu His Arg Ala Met Asn Leu His Val Gly Asn Leu Arg
 460 465 470

ctg ctc agc ggg ccg ctt gac cag gtc cgg gct gcc ctg ccc aca ccg 4950
 Leu Leu Ser Gly Pro Leu Asp Gln Val Arg Ala Ala Leu Pro Thr Pro
 475 480 485 490

gcc ctc tcc cca g gtgagcccca ccagacccca ttgggagact cgagctgggg 5003
 Ala Leu Ser Pro

gtttctctgg cctcaccgac cactgctgcc cacag ag gac aag gcc gtg ctg caa 5058
 Glu Asp Lys Ala Val Leu Gln
 495 500

aac cta aag cgc atc ctg gct aag gtg cag gag atg cgg gac cag cgc 5106
 Asn Leu Lys Arg Ile Leu Ala Lys Val Gln Glu Met Arg Asp Gln Arg
 505 510 515

gtg tcc ctg gag cag cag ctg cgt gag ctt atc cag aaa gat gac atc 5154
 Val Ser Leu Glu Gln Gln Leu Arg Glu Leu Ile Gln Lys Asp Asp Ile
 520 525 530

act gcc tcg ctg gtc acc aca gac cac tca gag atg aag gtgggctggg 5203
 Thr Ala Ser Leu Val Thr Thr Asp His Ser Glu Met Lys
 535 540 545

tgagcagggt agaggggctc tggctccggg cccaccctt aggagtcgag gccctgagtg 5263

tccgtccctg gccccaccc cttcctcag aag ttg ttc gag gag cag ctg aaa 5316

Lys Leu Phe Glu Glu Gln Leu Lys
550

aag tat gac cag ctg aag gtg tac ctg gag cag aac ctg gcc gcc cag 5364
Lys Tyr Asp Gln Leu Lys Val Tyr Leu Glu Gln Asn Leu Ala Ala Gln
555 560 565 570

gac cgt gtc ctc tgt gca ctg aca gag gcc aac gtg cag tac gca gcc 5412
Asp Arg Val Leu Cys Ala Leu Thr Glu Ala Asn Val Gln Tyr Ala Ala
575 580 585

gtg cgg cgg gta ctc agc gac ttg gac caa aa gtcagtgcc agtcctctgt 5464
Val Arg Arg Val Leu Ser Asp Leu Asp Gln Lys
590 595

cccttcccgg agccacctgg agcccagccc catggttcac ctggagctgg cccttctgcc 5524

caccag g tgg aac tcc acg ctg cag acc ctg gtg gcc tcg tat gaa gcc 5573
Trp Asn Ser Thr Leu Gln Thr Leu Val Ala Ser Tyr Glu Ala
600 605 610

tat gag gac ctg atg aag aag tcg cag gag ggc agg gac ttc tac gca 5621
Tyr Glu Asp Leu Met Lys Lys Ser Gln Glu Gly Arg Asp Phe Tyr Ala
615 620 625

gat ctg gag agc aag gtg gct gct ctg ctg gag cgc acg cag tcc acc 5669
Asp Leu Glu Ser Lys Val Ala Ala Leu Leu Glu Arg Thr Gln Ser Thr
630 635 640

tgc cag gcc cgc gag gct gcc cgc cag cag ctc ctg gac ag gtttgtgtgg 5720
Cys Gln Ala Arg Glu Ala Ala Arg Gln Gln Leu Leu Asp Arg
645 650 655

ccctggggct gtggtgcggt tcgggtccag acaggctggg gtgatgggag cctggcccca 5780

ctttttcett gcctgttgca cag g gag ctg aag aag aag ccg ccg cca cgg 5831
Glu Leu Lys Lys Lys Pro Pro Pro Arg

660	665	
ccc aca gcc cca aag ccg ctg ctg ccc cgc agg gag gag agt gag gca		5879
Pro Thr Ala Pro Lys Pro Leu Leu Pro Arg Arg Glu Glu Ser Glu Ala		
670	675	680
gtg gaa gca gga gac ccc cct gag gag ctg cgc agc ctc ccc cct gac		5927
Val Glu Ala Gly Asp Pro Pro Glu Glu Leu Arg Ser Leu Pro Pro Asp		
685	690	695
atg gtg gct ggc cca cga ctg cct gac acc ttc ctg gga agt gcc acc		5975
Met Val Ala Gly Pro Arg Leu Pro Asp Thr Phe Leu Gly Ser Ala Thr		
700	705	710
ccg ctc cac ttt cct ccc agc ccc ttc ccc agc tcc aca ggc cca gga		6023
Pro Leu His Phe Pro Pro Ser Pro Phe Pro Ser Ser Thr Gly Pro Gly		
715	720	725
		730
ccc cac tat ctc tca ggc ccc ttg ccc cct ggt acc tac tcg ggc ccc		6071
Pro His Tyr Leu Ser Gly Pro Leu Pro Pro Gly Thr Tyr Ser Gly Pro		
735	740	745
acc cag ctg ata cag ccc agg gcc cca ggg ccc cat gca atg ccc gta		6119
Thr Gln Leu Ile Gln Pro Arg Ala Pro Gly Pro His Ala Met Pro Val		
750	755	760
gca cct ggg cct gcc ctc tac cca gcc cct gca tac aca ccg gag ctg		6167
Ala Pro Gly Pro Ala Leu Tyr Pro Ala Pro Ala Tyr Thr Pro Glu Leu		
765	770	775
ggc ctt gtg ccc cga tcc tcc cca cag cat ggc gtg gtg agc agt ccc		6215
Gly Leu Val Pro Arg Ser Ser Pro Gln His Gly Val Val Ser Ser Pro		
780	785	790
tat gtg ggg gta ggg ccg gcc cca cca gtt gca ggt ctc ccc tcg gcc		6263
Tyr Val Gly Val Gly Pro Ala Pro Pro Val Ala Gly Leu Pro Ser Ala		
795	800	805
		810

cca cct cct caa ttc tca ggc ccc gag ttg gcc atg gcg gtt cgg cca 6311
 Pro Pro Pro Gln Phe Ser Gly Pro Glu Leu Ala Met Ala Val Arg Pro
 815 820 825

gcc acc acc aca gta gat agc atc cag gcg ccc atc ccc agc cac aca 6359
 Ala Thr Thr Thr Val Asp Ser Ile Gln Ala Pro Ile Pro Ser His Thr
 830 835 840

gcc cca cgg cca aac ccc acc cct gct cct ccc ccg ccc tgc ttc cct 6407
 Ala Pro Arg Pro Asn Pro Thr Pro Ala Pro Pro Pro Pro Cys Phe Pro
 845 850 855

gtg ccc cca ccg cag cca ctg ccc acg cct tac acc tac cct gca ggg 6455
 Val Pro Pro Pro Gln Pro Leu Pro Thr Pro Tyr Thr Tyr Pro Ala Gly
 860 865 870

gct aag caa ccc atc cca gca cag cac cac ttc tct tct ggg atc ccc 6503
 Ala Lys Gln Pro Ile Pro Ala Gln His His Phe Ser Ser Gly Ile Pro
 875 880 885 890

gca ggt ttt cca gcc cca agg att ggg ccc cag ccc cag ccc cat cct 6551
 Ala Gly Phe Pro Ala Pro Arg Ile Gly Pro Gln Pro Gln Pro His Pro
 895 900 905

cag ccc cat cct tca caa gcg ttt ggg cct cag ccc cca cag cag ccc 6599
 Gln Pro His Pro Ser Gln Ala Phe Gly Pro Gln Pro Pro Gln Gln Pro
 910 915 920

ctt cca ctc cag cat cca cat ctc ttc cca ccc cag gcc cca gga ctc 6647
 Leu Pro Leu Gln His Pro His Leu Phe Pro Pro Gln Ala Pro Gly Leu
 925 930 935

cta ccc cca caa tcc ccc tac ccc tat gcc cct cag cct ggg gtc ctg 6695
 Leu Pro Pro Gln Ser Pro Tyr Pro Tyr Ala Pro Gln Pro Gly Val Leu
 940 945 950

ggg cag ccg cca ccc ccc cta cac acc cag ctc tac cca ggt ccc gct 6743
 Gly Gln Pro Pro Pro Pro Leu His Thr Gln Leu Tyr Pro Gly Pro Ala
 955 960 965 970

caa gac cct ctg cca gcc cac tca ggg gct ctg cct ttc ccc agc cct 6791
 Gln Asp Pro Leu Pro Ala His Ser Gly Ala Leu Pro Phe Pro Ser Pro
 975 980 985

ggg ccc cct cag cct ccc cat ccc cca ctg gca tat ggt cct gcc cct 6839
 Gly Pro Pro Gln Pro Pro His Pro Pro Leu Ala Tyr Gly Pro Ala Pro
 990 995 1000

tct acc aga ccc atg ggc ccc cag gca gcc cct ctt acc att cga ggg 6887
 Ser Thr Arg Pro Met Gly Pro Gln Ala Ala Pro Leu Thr Ile Arg Gly
 1005 1010 1015

ccc tcg tct gct ggc cag tcc acc cct agt ccc cac ctg gtg cct tca 6935
 Pro Ser Ser Ala Gly Gln Ser Thr Pro Ser Pro His Leu Val Pro Ser
 1020 1025 1030

cct gcc cca tct cca ggg cct ggt ccg gta ccc cct cgc ccc cca gca 6983
 Pro Ala Pro Ser Pro Gly Pro Gly Pro Val Pro Pro Arg Pro Pro Ala
 1035 1040 1045 1050

gca gaa cca ccc cct tgc ctg cgc cga ggc gcc gca gct gca gac ctg 7031
 Ala Glu Pro Pro Pro Cys Leu Arg Arg Gly Ala Ala Ala Ala Asp Leu
 1055 1060 1065

ctc tcc tcc agc ccg gag agc cag cat ggc ggc act cag tct cct ggg 7079
 Leu Ser Ser Ser Pro Glu Ser Gln His Gly Gly Thr Gln Ser Pro Gly
 1070 1075 1080

ggt ggg cag ccc ctg ctg cag ccc acc aag gtg gat gca gct gag ggt 7127
 Gly Gly Gln Pro Leu Leu Gln Pro Thr Lys Val Asp Ala Ala Glu Gly
 1085 1090 1095

cgt cgg ccg cag gcc ctg cgg ctg att gag cgg gac ccc tat gag cat 7175

Arg Arg Pro Gln Ala Leu Arg Leu Ile Glu Arg Asp Pro Tyr Glu His	
1100	1105 1110
cct gag agg ctg cgg cag ttg cag cag gag ctg gag gcc ttt cgg ggt	7223
Pro Glu Arg Leu Arg Gln Leu Gln Gln Glu Leu Glu Ala Phe Arg Gly	
1115	1120 1125 1130
cag ctg ggg gat gtg gga gct ctg gac act gtc tgg cga gag ctg caa	7271
Gln Leu Gly Asp Val Gly Ala Leu Asp Thr Val Trp Arg Glu Leu Gln	
1135	1140 1145
gat gcg cag gaa cat gat gcc cga ggc cgt tcc atc gcc att gcc cgc	7319
Asp Ala Gln Glu His Asp Ala Arg Gly Arg Ser Ile Ala Ile Ala Arg	
1150	1155 1160
tgc tac tca ctg aag aac cgg cac cag gat gtc atg ccc tat gac agt	7367
Cys Tyr Ser Leu Lys Asn Arg His Gln Asp Val Met Pro Tyr Asp Ser	
1165	1170 1175
aac cgt gtg gtg ctg cgc tca ggc aag gat gac tac atc aat gcc agc	7415
Asn Arg Val Val Leu Arg Ser Gly Lys Asp Asp Tyr Ile Asn Ala Ser	
1180	1185 1190
tgc gtg gag ggg ctc tcc cca tac tgc ccc ccg cta gtg gca acc cag	7463
Cys Val Glu Gly Leu Ser Pro Tyr Cys Pro Pro Leu Val Ala Thr Gln	
1195	1200 1205 1210
gcc cca ctg cct ggc aca gct gct gac ttc tgg ctc atg gtc cat gag	7511
Ala Pro Leu Pro Gly Thr Ala Ala Asp Phe Trp Leu Met Val His Glu	
1215	1220 1225
cag aaa gtg tca gtc att gtc atg ctg gtt tct gag gct gag atg gag	7559
Gln Lys Val Ser Val Ile Val Met Leu Val Ser Glu Ala Glu Met Glu	
1230	1235 1240
aag gtgagaagag ggggtgggtgc ccccgaggca gtgtgggggtg gcagggcagg	7612
Lys	

ggatcctgga aaaccaggtc tgtcttggtc tatctgtccc tcag caa aaa gtg gca 7668
 Gln Lys Val Ala
 1245

cgc tac ttc ccc acc gag agg ggc cag ccc atg gtg cac ggt gcc ctg 7716
 Arg Tyr Phe Pro Thr Glu Arg Gly Gln Pro Met Val His Gly Ala Leu
 1250 1255 1260

agc ctg gca ttg agc agc gtc cgc agc acc gaa acc cat gtg gag cgc 7764
 Ser Leu Ala Leu Ser Ser Val Arg Ser Thr Glu Thr His Val Glu Arg
 1265 1270 1275

gtg ctg agc ctg cag ttc cga gac cag agc ctc aag cgc tct ctt gtg 7812
 Val Leu Ser Leu Gln Phe Arg Asp Gln Ser Leu Lys Arg Ser Leu Val
 1280 1285 1290 1295

cac ctg cac ttc ccc act tgg cct gag tt gtgagtcac tgcctctggat 7861
 His Leu His Phe Pro Thr Trp Pro Glu Leu
 1300 1305

ggtggttggg ggtctaagtg ctgtccagtc cttggtgctg ggagggatga gaggctcagg 7921

tcaggcctgg ctcataaggt cttcctggcc ccacctgtc ccacag a ggc ctg ccc 7977
 Gly Leu Pro

gac agc ccc agc aac ttg ctg cgc ttc atc cag gag gtg cac gca cat 8025
 Asp Ser Pro Ser Asn Leu Leu Arg Phe Ile Gln Glu Val His Ala His
 1310 1315 1320

tac ctg cat cag cgg ccg ctg cac acg ccc atc att gtg cac tgc ag 8072
 Tyr Leu His Gln Arg Pro Leu His Thr Pro Ile Ile Val His Cys Ser
 1325 1330 1335 1340

gtagagggtg ggcctgaggg tctctctctt atgggtcttt ggcctagcct cataccccgg 8132

cctcataacc cttcttggc acag c tct ggt gtg ggc cgc acg gga gcc ttt 8184

Ser Gly Val Gly Arg Thr Gly Ala Phe
1345

gca ctg ctc tat gca gct gtg cag gag gtg gag gct ggg aac gga atc 8232
Ala Leu Leu Tyr Ala Ala Val Gln Glu Val Glu Ala Gly Asn Gly Ile
1350 1355 1360 1365

cct gag ctg cct cag ctg gtg cgg cgc atg cgg cag cag aga aag cac 8280
Pro Glu Leu Pro Gln Leu Val Arg Arg Met Arg Gln Gln Arg Lys His
1370 1375 1380

atg ctg cag gag aag gtgatgatct gggcatatgg ggctgggatg ggccttctgt 8335
Met Leu Gln Glu Lys
1385

cccagggtga cgggccctg cccagctgac ctggccaaat gcacctgtgc ag ctg cac 8393
Leu His

ctc agg ttc tgc tat gag gca gtg gtg aga cac gtg gag cag gtc ctg 8441
Leu Arg Phe Cys Tyr Glu Ala Val Val Arg His Val Glu Gln Val Leu
1390 1395 1400

cag cgc cat ggt gtg cct cct cca tgc aaa ccc ttg gcc agt gca agc 8489
Gln Arg His Gly Val Pro Pro Pro Cys Lys Pro Leu Ala Ser Ala Ser
1405 1410 1415 1420

atc agc cag aag gtgaggaagg ttccgtggaa gctgctggga gagccacagc 8541
Ile Ser Gln Lys

cttggaatc cctctcctca ctcactctgt cttctcag aac cac ctt cct cag gac 8597
Asn His Leu Pro Gln Asp
1425 1430

tcc cag gac ctg gtc ctc ggt ggg gat gtg ccc atc agc tcc atc cag 8645
Ser Gln Asp Leu Val Leu Gly Gly Asp Val Pro Ile Ser Ser Ile Gln
1435 1440 1445

gcc acc att gcc aag ctc agc att cgg cct cct ggg ggg ttg gag tcc 8693
 Ala Thr Ile Ala Lys Leu Ser Ile Arg Pro Pro Gly Gly Leu Glu Ser
 1450 1455 1460

ccg gtt gcc agc ttg cca ggc cct gca gag ccc cca ggc ctc ccg cca 8741
 Pro Val Ala Ser Leu Pro Gly Pro Ala Glu Pro Pro Gly Leu Pro Pro
 1465 1470 1475

gcc agc ctc cca gag tct acc cca atc cca tct tcc tcc cca ccc ccc 8789
 Ala Ser Leu Pro Glu Ser Thr Pro Ile Pro Ser Ser Ser Pro Pro Pro
 1480 1485 1490

ctt tcc tcc cca cta cct gag gct ccc cag cct aag gag gag ccg cca 8837
 Leu Ser Ser Pro Leu Pro Glu Ala Pro Gln Pro Lys Glu Glu Pro Pro
 1495 1500 1505 1510

gtg cct gaa gcc ccc agc tcg ggg ccc ccc tcc tcc tcc ctg gaa ttg 8885
 Val Pro Glu Ala Pro Ser Ser Gly Pro Pro Ser Ser Ser Leu Glu Leu
 1515 1520 1525

ctg gcc tcc ttg acc cca gag gcc ttc tcc ctg gac agc tcc ctg cgg 8933
 Leu Ala Ser Leu Thr Pro Glu Ala Phe Ser Leu Asp Ser Ser Leu Arg
 1530 1535 1540

ggc aaa cag cgg atg agc aag cat aac ttt ctg cag gcc cat aac ggg 8981
 Gly Lys Gln Arg Met Ser Lys His Asn Phe Leu Gln Ala His Asn Gly
 1545 1550 1555

caa ggg ctg cgg gcc acc cgg ccc tct gac gac ccc ctc agc ctt ctg 9029
 Gln Gly Leu Arg Ala Thr Arg Pro Ser Asp Asp Pro Leu Ser Leu Leu
 1560 1565 1570

gat cca ctc tgg aca ctc aac aag acc tgaacaggtt ttgcctacct 9076
 Asp Pro Leu Trp Thr Leu Asn Lys Thr
 1575 1580

ggtccttaca ctacatcacc atcatctcat gccacactgc ccacaccag cagagcttct 9136

cagtgggcac agtctcttac tcccatttct gctgcctttg gccctgcctg gccagcctg 9196

cacccctgtg ggggtggaat gtactgcagg ctctgggtca ggttctgctc ctttatggga 9256

cccgacattt ttcagctctt tgctattgaa ataataaacc accctgttct gtg 9309

<210> 4

<211> 4022

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(3759)

<400> 4

atg gcc aag att gag gac aag aat gag gtc ctg gac cag ttc atg gat	48
Met Ala Lys Ile Glu Asp Lys Asn Glu Val Leu Asp Gln Phe Met Asp	
1 5 10 15	

tca atg cag ttg gat ccc gag acg gtg gac aac ctt gat gcc tac agc	96
Ser Met Gln Leu Asp Pro Glu Thr Val Asp Asn Leu Asp Ala Tyr Ser	
20 25 30	

cac atc cca ccc cag ctc atg gag aag tgc gcg gct ctc agc gtc cgg	144
His Ile Pro Pro Gln Leu Met Glu Lys Cys Ala Ala Leu Ser Val Arg	
35 40 45	

ccc gac act gtc agg aac ctt gta cag tcc atg caa gtg ctg tca ggt	192
Pro Asp Thr Val Arg Asn Leu Val Gln Ser Met Gln Val Leu Ser Gly	
50 55 60	

gtg ttc acg gat gtg gag gct tcc ctg aag gac atc aga gat ctg ttg	240
Val Phe Thr Asp Val Glu Ala Ser Leu Lys Asp Ile Arg Asp Leu Leu	
65 70 75 80	

gag gag gat gag ctg cta gag cag aag ttt cag gag gcg gtg ggc cag	288
Glu Glu Asp Glu Leu Leu Glu Gln Lys Phe Gln Glu Ala Val Gly Gln	
85 90 95	

gca ggg gcc atc tcc atc acc tcc aag gct gag ctg gca gag gtg agg	336
Ala Gly Ala Ile Ser Ile Thr Ser Lys Ala Glu Leu Ala Glu Val Arg	
100 105 110	
cga gaa tgg gcc aag tac atg gaa gtc cat gag aag gcc tcc ttc acc	384
Arg Glu Trp Ala Lys Tyr Met Glu Val His Glu Lys Ala Ser Phe Thr	
115 120 125	
aac agt gag ctg cac cgt gcc atg aac ctg cac gtc ggc aac ctg cgc	432
Asn Ser Glu Leu His Arg Ala Met Asn Leu His Val Gly Asn Leu Arg	
130 135 140	
ctg ctc agc ggg ccg ctt gac cag gtc cgg gct gcc ctg ccc aca ccg	480
Leu Leu Ser Gly Pro Leu Asp Gln Val Arg Ala Ala Leu Pro Thr Pro	
145 150 155 160	
gcc ctc tcc cca gag gac aag gcc gtg ctg caa aac cta aag cgc atc	528
Ala Leu Ser Pro Glu Asp Lys Ala Val Leu Gln Asn Leu Lys Arg Ile	
165 170 175	
ctg gct aag gtg cag gag atg cgg gac cag cgc gtg tcc ctg gag cag	576
Leu Ala Lys Val Gln Glu Met Arg Asp Gln Arg Val Ser Leu Glu Gln	
180 185 190	
cag ctg cgt gag ctt atc cag aaa gat gac atc act gcc tcg ctg gtc	624
Gln Leu Arg Glu Leu Ile Gln Lys Asp Asp Ile Thr Ala Ser Leu Val	
195 200 205	
acc aca gac cac tca gag atg aag aag ttg ttc gag gag cag ctg aaa	672
Thr Thr Asp His Ser Glu Met Lys Lys Leu Phe Glu Glu Gln Leu Lys	
210 215 220	
aag tat gac cag ctg aag gtg tac ctg gag cag aac ctg gcc gcc cag	720
Lys Tyr Asp Gln Leu Lys Val Tyr Leu Glu Gln Asn Leu Ala Ala Gln	
225 230 235 240	
gac cgt gtc ctc tgt gca ctg aca gag gcc aac gtg cag tac gca gcc	768
Asp Arg Val Leu Cys Ala Leu Thr Glu Ala Asn Val Gln Tyr Ala Ala	
245 250 255	
gtg cgg cgg gta ctc agc gac ttg gac caa aag tgg aac tcc acg ctg	816

Val Arg Arg Val Leu Ser Asp Leu Asp Gln Lys Trp Asn Ser Thr Leu	
260 265 270	
cag acc ctg gtg gcc tcg tat gaa gcc tat gag gac ctg atg aag aag	864
Gln Thr Leu Val Ala Ser Tyr Glu Ala Tyr Glu Asp Leu Met Lys Lys	
275 280 285	
tcg cag gag ggc agg gac ttc tac gca gat ctg gag agc aag gtg gct	912
Ser Gln Glu Gly Arg Asp Phe Tyr Ala Asp Leu Glu Ser Lys Val Ala	
290 295 300	
gct ctg ctg gag cgc acg cag tcc acc tgc cag gcc cgc gag gct gcc	960
Ala Leu Leu Glu Arg Thr Gln Ser Thr Cys Gln Ala Arg Glu Ala Ala	
305 310 315 320	
cgc cag cag ctc ctg gac agg gag ctg aag aag aag ccg ccg cca cgg	1008
Arg Gln Gln Leu Leu Asp Arg Glu Leu Lys Lys Lys Pro Pro Pro Arg	
325 330 335	
ccc aca gcc cca aag ccg ctg ctg ccc cgc agg gag gag agt gag gca	1056
Pro Thr Ala Pro Lys Pro Leu Leu Pro Arg Arg Glu Glu Ser Glu Ala	
340 345 350	
gtg gaa gca gga gac ccc cct gag gag ctg cgc agc ctc ccc cct gac	1104
Val Glu Ala Gly Asp Pro Pro Glu Glu Leu Arg Ser Leu Pro Pro Asp	
355 360 365	
atg gtg gct ggc cca cga ctg cct gac acc ttc ctg gga agt gcc acc	1152
Met Val Ala Gly Pro Arg Leu Pro Asp Thr Phe Leu Gly Ser Ala Thr	
370 375 380	
ccg ctc cac ttt cct ccc agc ccc ttc ccc agc tcc aca ggc cca gga	1200
Pro Leu His Phe Pro Pro Ser Pro Phe Pro Ser Ser Thr Gly Pro Gly	
385 390 395 400	
ccc cac tat ctc tca ggc ccc ttg ccc cct ggt acc tac tcg ggc ccc	1248
Pro His Tyr Leu Ser Gly Pro Leu Pro Pro Gly Thr Tyr Ser Gly Pro	
405 410 415	
acc cag ctg ata cag ccc agg gcc cca ggg ccc cat gca atg ccc gta	1296
Thr Gln Leu Ile Gln Pro Arg Ala Pro Gly Pro His Ala Met Pro Val	
420 425 430	

gca cct ggg cct gcc ctc tac cca gcc cct gca tac aca ccg gag ctg	1344
Ala Pro Gly Pro Ala Leu Tyr Pro Ala Pro Ala Tyr Thr Pro Glu Leu	
435 440 445	
ggc ctt gtg ccc cga tcc tcc cca cag cat ggc gtg gtg agc agt ccc	1392
Gly Leu Val Pro Arg Ser Ser Pro Gln His Gly Val Val Ser Ser Pro	
450 455 460	
tat gtg ggg gta ggg ccg gcc cca cca gtt gca ggt ctc ccc tcg gcc	1440
Tyr Val Gly Val Gly Pro Ala Pro Pro Val Ala Gly Leu Pro Ser Ala	
465 470 475 480	
cca cct cct caa ttc tca ggc ccc gag ttg gcc atg gcg gtt cgg cca	1488
Pro Pro Pro Gln Phe Ser Gly Pro Glu Leu Ala Met Ala Val Arg Pro	
485 490 495	
gcc acc acc aca gta gat agc atc cag gcg ccc atc ccc agc cac aca	1536
Ala Thr Thr Thr Val Asp Ser Ile Gln Ala Pro Ile Pro Ser His Thr	
500 505 510	
gcc cca cgg cca aac ccc acc cct gct cct ccc ccg ccc tgc ttc cct	1584
Ala Pro Arg Pro Asn Pro Thr Pro Ala Pro Pro Pro Pro Cys Phe Pro	
515 520 525	
gtg ccc cca ccg cag cca ctg ccc acg cct tac acc tac cct gca ggg	1632
Val Pro Pro Pro Gln Pro Leu Pro Thr Pro Tyr Thr Tyr Pro Ala Gly	
530 535 540	
gct aag caa ccc atc cca gca cag cac cac ttc tct tct ggg atc ccc	1680
Ala Lys Gln Pro Ile Pro Ala Gln His His Phe Ser Ser Gly Ile Pro	
545 550 555 560	
gca ggt ttt cca gcc cca agg att ggg ccc cag ccc cag ccc cat cct	1728
Ala Gly Phe Pro Ala Pro Arg Ile Gly Pro Gln Pro Gln Pro His Pro	
565 570 575	
cag ccc cat cct tca caa gcg ttt ggg cct cag ccc cca cag cag ccc	1776
Gln Pro His Pro Ser Gln Ala Phe Gly Pro Gln Pro Pro Gln Gln Pro	
580 585 590	
ctt cca ctc cag cat cca cat ctc ttc cca ccc cag gcc cca gga ctc	1824

Leu Pro Leu Gln His Pro His Leu Phe Pro Pro Gln Ala Pro Gly Leu	
595	600 605
cta ccc cca caa tcc ccc tac ccc tat gcc cct cag cct ggg gtc ctg	1872
Leu Pro Pro Gln Ser Pro Tyr Pro Tyr Ala Pro Gln Pro Gly Val Leu	
610	615 620
ggg cag ccg cca ccc ccc cta cac acc cag ctc tac cca ggt ccc gct	1920
Gly Gln Pro Pro Pro Pro Leu His Thr Gln Leu Tyr Pro Gly Pro Ala	
625	630 635 640
caa gac cct ctg cca gcc cac tca ggg gct ctg cct ttc ccc agc cct	1968
Gln Asp Pro Leu Pro Ala His Ser Gly Ala Leu Pro Phe Pro Ser Pro	
	645 650 655
ggg ccc cct cag cct ccc cat ccc cca ctg gca tat ggt cct gcc cct	2016
Gly Pro Pro Gln Pro Pro His Pro Pro Leu Ala Tyr Gly Pro Ala Pro	
	660 665 670
tct acc aga ccc atg ggc ccc cag gca gcc cct ctt acc att cga ggg	2064
Ser Thr Arg Pro Met Gly Pro Gln Ala Ala Pro Leu Thr Ile Arg Gly	
	675 680 685
ccc tcg tct gct ggc cag tcc acc cct agt ccc cac ctg gtg cct tca	2112
Pro Ser Ser Ala Gly Gln Ser Thr Pro Ser Pro His Leu Val Pro Ser	
	690 695 700
cct gcc cca tct cca ggg cct ggt ccg gta ccc cct cgc ccc cca gca	2160
Pro Ala Pro Ser Pro Gly Pro Gly Pro Val Pro Pro Arg Pro Pro Ala	
705	710 715 720
gca gaa cca ccc cct tgc ctg cgc cga ggc gcc gca gct gca gac ctg	2208
Ala Glu Pro Pro Pro Cys Leu Arg Arg Gly Ala Ala Ala Ala Asp Leu	
	725 730 735
ctc tcc tcc agc ccg gag agc cag cat ggc ggc act cag tct cct ggg	2256
Leu Ser Ser Ser Pro Glu Ser Gln His Gly Gly Thr Gln Ser Pro Gly	
	740 745 750
ggt ggg cag ccc ctg ctg cag ccc acc aag gtg gat gca gct gag ggt	2304
Gly Gly Gln Pro Leu Leu Gln Pro Thr Lys Val Asp Ala Ala Glu Gly	
	755 760 765

cgt cgg ccg cag gcc ctg cgg ctg att gag cgg gac ccc tat gag cat	2352
Arg Arg Pro Gln Ala Leu Arg Leu Ile Glu Arg Asp Pro Tyr Glu His	
770 775 780	
cct gag agg ctg cgg cag ttg cag cag gag ctg gag gcc ttt cgg ggt	2400
Pro Glu Arg Leu Arg Gln Leu Gln Gln Glu Leu Glu Ala Phe Arg Gly	
785 790 795 800	
cag ctg ggg gat gtg gga gct ctg gac act gtc tgg cga gag ctg caa	2448
Gln Leu Gly Asp Val Gly Ala Leu Asp Thr Val Trp Arg Glu Leu Gln	
805 810 815	
gat gcg cag gaa cat gat gcc cga ggc cgt tcc atc gcc att gcc cgc	2496
Asp Ala Gln Glu His Asp Ala Arg Gly Arg Ser Ile Ala Ile Ala Arg	
820 825 830	
tgc tac tca ctg aag aac cgg cac cag gat gtc atg ccc tat gac agt	2544
Cys Tyr Ser Leu Lys Asn Arg His Gln Asp Val Met Pro Tyr Asp Ser	
835 840 845	
aac cgt gtg gtg ctg cgc tca ggc aag gat gac tac atc aat gcc agc	2592
Asn Arg Val Val Leu Arg Ser Gly Lys Asp Asp Tyr Ile Asn Ala Ser	
850 855 860	
tgc gtg gag ggg ctc tcc cca tac tgc ccc ccg cta gtg gca acc cag	2640
Cys Val Glu Gly Leu Ser Pro Tyr Cys Pro Pro Leu Val Ala Thr Gln	
865 870 875 880	
gcc cca ctg cct ggc aca gct gct gac ttc tgg ctc atg gtc cat gag	2688
Ala Pro Leu Pro Gly Thr Ala Ala Asp Phe Trp Leu Met Val His Glu	
885 890 895	
cag aaa gtg tca gtc att gtc atg ctg gtt tct gag gct gag atg gag	2736
Gln Lys Val Ser Val Ile Val Met Leu Val Ser Glu Ala Glu Met Glu	
900 905 910	
aag caa aaa gtg gca cgc tac ttc ccc acc gag agg ggc cag ccc atg	2784
Lys Gln Lys Val Ala Arg Tyr Phe Pro Thr Glu Arg Gly Gln Pro Met	
915 920 925	
gtg cac ggt gcc ctg agc ctg gca ttg agc agc gtc cgc agc acc gaa	2832

Val His Gly Ala Leu Ser Leu Ala Leu Ser Ser Val Arg Ser Thr Glu	
930 935 940	
acc cat gtg gag cgc gtg ctg agc ctg cag ttc cga gac cag agc ctc	2880
Thr His Val Glu Arg Val Leu Ser Leu Gln Phe Arg Asp Gln Ser Leu	
945 950 955 960	
aag cgc tct ctt gtg cac ctg cac ttc ccc act tgg cct gag tta ggc	2928
Lys Arg Ser Leu Val His Leu His Phe Pro Thr Trp Pro Glu Leu Gly	
965 970 975	
ctg ccc gac agc ccc agc aac ttg ctg cgc ttc atc cag gag gtg cac	2976
Leu Pro Asp Ser Pro Ser Asn Leu Leu Arg Phe Ile Gln Glu Val His	
980 985 990	
gca cat tac ctg cat cag cgg cgg ctg cac acg ccc atc att gtg cac	3024
Ala His Tyr Leu His Gln Arg Pro Leu His Thr Pro Ile Ile Val His	
995 1000 1005	
tgc agc tct ggt gtg ggc cgc acg gga gcc ttt gca ctg ctc tat gca	3072
Cys Ser Ser Gly Val Gly Arg Thr Gly Ala Phe Ala Leu Leu Tyr Ala	
1010 1015 1020	
gct gtg cag gag gtg gag gct ggg aac gga atc cct gag ctg cct cag	3120
Ala Val Gln Glu Val Glu Ala Gly Asn Gly Ile Pro Glu Leu Pro Gln	
1025 1030 1035 1040	
ctg gtg cgg cgc atg cgg cag cag aga aag cac atg ctg cag gag aag	3168
Leu Val Arg Arg Met Arg Gln Gln Arg Lys His Met Leu Gln Glu Lys	
1045 1050 1055	
ctg cac ctc agg ttc tgc tat gag gca gtg gtg aga cac gtg gag cag	3216
Leu His Leu Arg Phe Cys Tyr Glu Ala Val Val Arg His Val Glu Gln	
1060 1065 1070	
gtc ctg cag cgc cat ggt gtg cct cct cca tgc aaa ccc ttg gcc agt	3264
Val Leu Gln Arg His Gly Val Pro Pro Pro Cys Lys Pro Leu Ala Ser	
1075 1080 1085	
gca agc atc agc cag aag aac cac ctt cct cag gac tcc cag gac ctg	3312
Ala Ser Ile Ser Gln Lys Asn His Leu Pro Gln Asp Ser Gln Asp Leu	
1090 1095 1100	

gtc ctc ggt ggg gat gtg ccc atc agc tcc atc cag gcc acc att gcc 3360
 Val Leu Gly Gly Asp Val Pro Ile Ser Ser Ile Gln Ala Thr Ile Ala
 1105 1110 1115 1120

aag ctc agc att cgg cct cct ggg ggg ttg gag tcc ccg gtt gcc agc 3408
 Lys Leu Ser Ile Arg Pro Pro Gly Gly Leu Glu Ser Pro Val Ala Ser
 1125 1130 1135

ttg cca ggc cct gca gag ccc cca ggc ctc ccg cca gcc agc etc cca 3456
 Leu Pro Gly Pro Ala Glu Pro Pro Gly Leu Pro Pro Ala Ser Leu Pro
 1140 1145 1150

gag tct acc cca atc cca tct tcc tcc cca ccc ccc ctt tcc tcc cca 3504
 Glu Ser Thr Pro Ile Pro Ser Ser Ser Pro Pro Pro Leu Ser Ser Pro
 1155 1160 1165

cta cct gag gct ccc cag cct aag gag gag ccg cca gtg cct gaa gcc 3552
 Leu Pro Glu Ala Pro Gln Pro Lys Glu Glu Pro Pro Val Pro Glu Ala
 1170 1175 1180

ccc agc tcg ggg ccc ccc tcc tcc tcc ctg gaa ttg ctg gcc tcc ttg 3600
 Pro Ser Ser Gly Pro Pro Ser Ser Ser Leu Glu Leu Leu Ala Ser Leu
 1185 1190 1195 1200

acc cca gag gcc ttc tcc ctg gac agc tcc ctg cgg ggc aaa cag cgg 3648
 Thr Pro Glu Ala Phe Ser Leu Asp Ser Ser Leu Arg Gly Lys Gln Arg
 1205 1210 1215

atg agc aag cat aac ttt ctg cag gcc cat aac ggg caa ggg ctg cgg 3696
 Met Ser Lys His Asn Phe Leu Gln Ala His Asn Gly Gln Gly Leu Arg
 1220 1225 1230

gcc acc cgg ccc tct gac gac ccc ctc agc ctt ctg gat cca ctc tgg 3744
 Ala Thr Arg Pro Ser Asp Asp Pro Leu Ser Leu Leu Asp Pro Leu Trp
 1235 1240 1245

aca ctc aac aag acc tgaacaggtt ttgcctacct ggtccttaca ctacatcatc 3799
 Thr Leu Asn Lys Thr
 1250

atcatctcat gccacctgc ccacaccag cagagcttct cagtgggcac agtctcttac 3859

tcccatttct gctgcctttg gccctgcctg gccagcctg caccctgtg gggtggaat 3919
gtactgcagg ctctgggtca ggttctgctc ctttatggga cccgacattt ttcagctctt 3979
tgctattgaa ataataaacc accctgttct gtgaaaaaaaa aaa 4022

<210> 5
<211> 5434
<212> DNA
<213> Homo sapiens

<220>
<221> exon
<222> (75)..(140)

<220>
<221> exon
<222> (530)..(675)

<220>
<221> exon
<222> (961)..(1272)

<220>
<221> exon
<222> (1348)..(1654)

<220>
<221> exon
<222> (1741)..(3679)

<220>
<221> exon
<222> (3774)..(3958)

<220>
<221> exon
<222> (4085)..(4189)

<220>

<221> exon
<222> (4274)..(4412)

<220>
<221> exon
<222> (4505)..(4618)

<220>
<221> exon
<222> (4697)..(5436)

<220>
<221> intron
<222> (1)..(74)
<220>
<221> intron
<222> (141)..(529)

<220>
<221> intron
<222> (676)..(960)

<220>
<221> intron
<222> (1273)..(1347)

<220>
<221> intron
<222> (1655)..(1740)

<220>
<221> intron
<222> (3680)..(3773)

<220>
<221> intron
<222> (3959)..(4084)

<220>
<221> intron
<222> (4190)..(4273)

<220>

<221> intron

<222> (4413)..(4504)

<220>

<221> intron

<222> (4619)..(4696)

<400> 5

gtggagggtg gcacagaggg aggtgggggtg tcttgagatg tgggtcttca gcaaagtctc 60

tgtcgttct gcag t gag gag aag gcc aag ctg ctc cgg gag atg atg 108

Ser Glu Glu Lys Ala Lys Leu Leu Arg Glu Met Met

1

5

10

gcc aag att gag gac aag aat gag gtc ctg ga gtgagtgtg gacttgggca 160

Ala Lys Ile Glu Asp Lys Asn Glu Val Leu Asp

15

20

gggaggcgga ggcaggcagc acttcccggg cctctggggg ccccagggt gcctatgctg 220

ggagaggaat gaaatgtcca ttccaaacag gtttcccaat gctgccttcc cgcccggggt 280

ggtggggcta gtgtgtaagg caggagtcac gtcttgggag gaggaggtgc cttctgttcc 340

actgtttcca gcagtgcctt gggcatgttc tgtgagacca ggccagacct ggtagtaggg 400

gtccgaggtc acaatttgct ctctgctgag acctcagatt gagtgggtgag gcttgcctta 460

gcggctcctt tgacatggtc agagttggat cagcaccag caccacctg gccctgttgc 520

tccccacag c cag ttc atg gat tca atg cag ttg gat ccc gag acg gtg 569

Gln Phe Met Asp Ser Met Gln Leu Asp Pro Glu Thr Val

25

30

35

gac aac ctt gat gcc tac agc cac atc cca ccc cag ctc atg gag aag 617

Asp Asn Leu Asp Ala Tyr Ser His Ile Pro Pro Gln Leu Met Glu Lys

40

45

50

tgc gcg gct ctc agc gtc cgg ccc gac act gtc agg aac ctt gta cag 665

Cys Ala Ala Leu Ser Val Arg Pro Asp Thr Val Arg Asn Leu Val Gln

55

60

65

tcc atg caa g gtgagtaagg ggcagagcaa gcaggtggaa gggagtgtgg 715
 Ser Met Gln
 70

aggtcatcta ctgtggcctc ctccgtgtcc ctggtcactg aggatggaga ctgcaccct 775
 ctaggccttg gcttgggcac ccacaccac tcctctgaat cagcatacct cttgcaccct 835
 gctcagtgtg cgctgggcct cacttaagcc ctgacctgag ggggcggttc tgtctcttgg 895
 gggagggggc catgggtgcc cggtcagcct gcctcagggg ctgcctgtac aatccacaac 955

cccag tg ctg tca ggt gtg ttc acg gat gtg gag gct tcc ctg aag gac 1004
 Val Leu Ser Gly Val Phe Thr Asp Val Glu Ala Ser Leu Lys Asp
 75 80 85

atc aga gat ctg ttg gag gag gat gag ctg cta gag cag aag ttt cag 1052
 Ile Arg Asp Leu Leu Glu Glu Asp Glu Leu Leu Glu Gln Lys Phe Gln
 90 95 100

gag gcg gtg ggc cag gca ggg gcc atc tcc atc acc tcc aag gct gag 1100
 Glu Ala Val Gly Gln Ala Gly Ala Ile Ser Ile Thr Ser Lys Ala Glu
 105 110 115

ctg gca gag gtg agg cga gaa tgg gcc aag tac atg gaa gtc cat gag 1148
 Leu Ala Glu Val Arg Arg Glu Trp Ala Lys Tyr Met Glu Val His Glu
 120 125 130

aag gcc tcc ttc acc aac agt gag ctg cac cgt gcc atg aac ctg cac 1196
 Lys Ala Ser Phe Thr Asn Ser Glu Leu His Arg Ala Met Asn Leu His
 135 140 145 150

gtc ggc aac ctg cgc ctg ctc agc ggg ccg ctt gac cag gtc cgg gct 1244
 Val Gly Asn Leu Arg Leu Leu Ser Gly Pro Leu Asp Gln Val Arg Ala
 155 160 165

gcc ctg ccc aca ccg gcc ctc tcc cca g gtgagcccca ccagacccca 1292
 Ala Leu Pro Thr Pro Ala Leu Ser Pro
 170 175

ttgggagact cgagctgggg gtttctctgg cctcaccgac cactgctgcc cacag ag 1349

Glu

gac aag gcc gtg ctg caa aac cta aag cgc atc ctg gct aag gtg cag 1397
 Asp Lys Ala Val Leu Gln Asn Leu Lys Arg Ile Leu Ala Lys Val Gln
 180 185 190

gag atg cgg gac cag cgc gtg tcc ctg gag cag cag ctg cgt gag ctt 1445
 Glu Met Arg Asp Gln Arg Val Ser Leu Glu Gln Gln Leu Arg Glu Leu
 195 200 205

atc cag aaa gat gac atc act gcc tcg ctg gtc acc aca gac cac tca 1493
 Ile Gln Lys Asp Asp Ile Thr Ala Ser Leu Val Thr Thr Asp His Ser
 210 215 220

gag atg aag aag ttg ttc gag gag cag ctg aaa aag tat gac cag ctg 1541
 Glu Met Lys Lys Leu Phe Glu Glu Gln Leu Lys Lys Tyr Asp Gln Leu
 225 230 235 240

aag gtg tac ctg gag cag aac ctg gcc gcc cag gac cgt gtc ctc tgt 1589
 Lys Val Tyr Leu Glu Gln Asn Leu Ala Ala Gln Asp Arg Val Leu Cys
 245 250 255

gca ctg aca gag gcc aac gtg cag tac gca gcc gtg cgg cgg gta ctc 1637
 Ala Leu Thr Glu Ala Asn Val Gln Tyr Ala Ala Val Arg Arg Val Leu
 260 265 270

agc gac ttg gac caa aa gtcagtcccc agtcctctgt cctttccgg 1684
 Ser Asp Leu Asp Gln Lys
 275

agccacctgg agcccagccc catggttcac ctggagctgg cccttctgcc caccag g 1741

tgg aac tcc acg ctg cag acc ctg gtg gcc tcg tat gaa gcc tat gag 1789
 Trp Asn Ser Thr Leu Gln Thr Leu Val Ala Ser Tyr Glu Ala Tyr Glu
 280 285 290

gac ctg atg aag aag tcg cag gag ggc agg gac ttc tac gca gat ctg 1837
 Asp Leu Met Lys Lys Ser Gln Glu Gly Arg Asp Phe Tyr Ala Asp Leu
 295 300 305 310

gag agc aag gtg gct gct ctg ctg gag cgc acg cag tcc acc tgc cag 1885

Glu Ser Lys Val Ala Ala Leu Leu Glu Arg Thr Gln Ser Thr Cys Gln	
315 320 325	
gcc cgc gag gct gcc cgc cag cag ctc ctg gac agg gag ctg aag aag	1933
Ala Arg Glu Ala Ala Arg Gln Gln Leu Leu Asp Arg Glu Leu Lys Lys	
330 335 340	
aag ccg ccg cca cgg ccc aca gcc cca aag ccg ctg ctg ccc cgc agg	1981
Lys Pro Pro Pro Arg Pro Thr Ala Pro Lys Pro Leu Leu Pro Arg Arg	
345 350 355	
gag gag agt gag gca gtg gaa gca gga gac ccc cct gag gag ctg cgc	2029
Glu Glu Ser Glu Ala Val Glu Ala Gly Asp Pro Pro Glu Glu Leu Arg	
360 365 370	
agc ctc ccc cct gac atg gtg gct ggc cca cga ctg cct gac acc ttc	2077
Ser Leu Pro Pro Asp Met Val Ala Gly Pro Arg Leu Pro Asp Thr Phe	
375 380 385 390	
ctg gga agt gcc acc ccg ctc cac ttt cct ccc agc ccc ttc ccc agc	2125
Leu Gly Ser Ala Thr Pro Leu His Phe Pro Pro Ser Pro Phe Pro Ser	
395 400 405	
tcc aca ggc cca gga ccc cac tat ctc tca ggc ccc ttg ccc cct ggt	2173
Ser Thr Gly Pro Gly Pro His Tyr Leu Ser Gly Pro Leu Pro Pro Gly	
410 415 420	
acc tac tcg ggc ccc acc cag ctg ata cag ccc agg gcc cca ggg ccc	2221
Thr Tyr Ser Gly Pro Thr Gln Leu Ile Gln Pro Arg Ala Pro Gly Pro	
425 430 435	
cat gca atg ccc gta gca cct ggg cct gcc ctc tac cca gcc cct gca	2269
His Ala Met Pro Val Ala Pro Gly Pro Ala Leu Tyr Pro Ala Pro Ala	
440 445 450	
tac aca ccg gag ctg ggc ctt gtg ccc cga tcc tcc cca cag cat ggc	2317
Tyr Thr Pro Glu Leu Gly Leu Val Pro Arg Ser Ser Pro Gln His Gly	
455 460 465 470	
gtg gtg agc agt ccc tat gtg ggg gta ggg ccg gcc cca cca gtt gca	2365
Val Val Ser Ser Pro Tyr Val Gly Val Gly Pro Ala Pro Pro Val Ala	
475 480 485	

ggt ctc ccc tcg gcc cca cct cct caa ttc tca ggc ccc gag ttg gcc 2413
 Gly Leu Pro Ser Ala Pro Pro Pro Gln Phe Ser Gly Pro Glu Leu Ala
 490 495 500

atg gcg gtt cgg cca gcc acc acc aca gta gat agc atc cag gcg ccc 2461
 Met Ala Val Arg Pro Ala Thr Thr Thr Val Asp Ser Ile Gln Ala Pro
 455 510 515

atc ccc agc cac aca gcc cca cgg cca aac ccc acc cct gct cct ccc 2509
 Ile Pro Ser His Thr Ala Pro Arg Pro Asn Pro Thr Pro Ala Pro Pro
 520 525 530

ccg ccc tgc ttc cct gtg ccc cca ccg cag cca ctg ccc acg cct tac 2557
 Pro Pro Cys Phe Pro Val Pro Pro Pro Gln Pro Leu Pro Thr Pro Tyr
 535 540 545 550

acc tac cct gca ggg gct aag caa ccc atc cca gca cag cac cac ttc 2605
 Thr Tyr Pro Ala Gly Ala Lys Gln Pro Ile Pro Ala Gln His His Phe
 555 560 565

tct tct ggg atc ccc gca ggt ttt cca gcc cca agg att ggg ccc cag 2653
 Ser Ser Gly Ile Pro Ala Gly Phe Pro Ala Pro Arg Ile Gly Pro Gln
 570 575 580

ccc cag ccc cat cct cag ccc cat cct tca caa gcg ttt ggg cct cag 2701
 Pro Gln Pro His Pro Gln Pro His Pro Ser Gln Ala Phe Gly Pro Gln
 585 590 595

ccc cca cag cag ccc ctt cca ctc cag cat cca cat ctc ttc cca ccc 2749
 Pro Pro Gln Gln Pro Leu Pro Leu Gln His Pro His Leu Phe Pro Pro
 600 605 610

cag gcc cca gga ctc cta ccc cca caa tcc ccc tac ccc tat gcc cct 2797
 Gln Ala Pro Gly Leu Leu Pro Pro Gln Ser Pro Tyr Pro Tyr Ala Pro
 615 620 625 630

cag cct ggg gtc ctg ggg cag ccg cca ccc ccc cta cac acc cag ctc 2845
 Gln Pro Gly Val Leu Gly Gln Pro Pro Pro Pro Leu His Thr Gln Leu
 635 640 645

tac cca ggt ccc gct caa gac cct ctg cca gcc cac tca ggg gct ctg 2893

Tyr Pro Gly Pro Ala Gln Asp Pro Leu Pro Ala His Ser Gly Ala Leu	
650	655 660
cct ttc ccc agc cct ggg ccc cct cag cct ccc cat ccc cca ctg gca	2941
Pro Phe Pro Ser Pro Gly Pro Pro Gln Pro Pro His Pro Pro Leu Ala	
665	670 675
tat ggt cct gcc cct tct acc aga ccc atg ggc ccc cag gca gcc cct	2989
Tyr Gly Pro Ala Pro Ser Thr Arg Pro Met Gly Pro Gln Ala Ala Pro	
680	685 690
ctt acc att cga ggg ccc tcg tct gct ggc cag tcc acc cct agt ccc	3037
Leu Thr Ile Arg Gly Pro Ser Ser Ala Gly Gln Ser Thr Pro Ser Pro	
695	700 705 710
cac ctg gtg cct tca cct gcc cca tct cca ggg cct ggt ccg gta ccc	3085
His Leu Val Pro Ser Pro Ala Pro Ser Pro Gly Pro Gly Pro Val Pro	
715	720 725
cct cgc ccc cca gca gca gaa cca ccc cct tgc ctg cgc cga ggc gcc	3133
Pro Arg Pro Pro Ala Ala Glu Pro Pro Pro Cys Leu Arg Arg Gly Ala	
730	735 740
gca gct gca gac ctg ctc tcc tcc agc ccg gag agc cag cat ggc ggc	3181
Ala Ala Ala Asp Leu Leu Ser Ser Ser Pro Glu Ser Gln His Gly Gly	
745	750 755
act cag tct cct ggg ggt ggg cag ccc ctg ctg cag ccc acc aag gtg	3229
Thr Gln Ser Pro Gly Gly Gly Gln Pro Leu Leu Gln Pro Thr Lys Val	
760	765 770
gat gca gct gag ggt cgt cgg ccg cag gcc ctg cgg ctg att gag cgg	3277
Asp Ala Ala Glu Gly Arg Arg Pro Gln Ala Leu Arg Leu Ile Glu Arg	
775	780 785 790
gac ccc tat gag cat cct gag agg ctg cgg cag ttg cag cag gag ctg	3325
Asp Pro Tyr Glu His Pro Glu Arg Leu Arg Gln Leu Gln Gln Glu Leu	
795	800 805
gag gcc ttt cgg ggt cag ctg ggg gat gtg gga gct ctg gac act gtc	3373
Glu Ala Phe Arg Gly Gln Leu Gly Asp Val Gly Ala Leu Asp Thr Val	
810	815 820

tgg cga gag ctg caa gat gcg cag gaa cat gat gcc cga ggc cgt tcc 3421
 Trp Arg Glu Leu Gln Asp Ala Gln Glu His Asp Ala Arg Gly Arg Ser
 825 830 835

atc gcc att gcc cgc tgc tac tca ctg aag aac cgg cac cag gat gtc 3469
 Ile Ala Ile Ala Arg Cys Tyr Ser Leu Lys Asn Arg His Gln Asp Val
 840 845 850

atg ccc tat gac agt aac cgt gtg gtg ctg cgc tca ggc aag gat gac 3517
 Met Pro Tyr Asp Ser Asn Arg Val Val Leu Arg Ser Gly Lys Asp Asp
 855 860 865 870

tac atc aat gcc agc tgc gtg gag ggg ctc tcc cca tac tgc ccc ccg 3565
 Tyr Ile Asn Ala Ser Cys Val Glu Gly Leu Ser Pro Tyr Cys Pro Pro
 875 880 885

cta gtg gca acc cag gcc cca ctg cct ggc aca gct gct gac ttc tgg 3613
 Leu Val Ala Thr Gln Ala Pro Leu Pro Gly Thr Ala Ala Asp Phe Trp
 890 895 900

ctc atg gtc cat gag cag aaa gtg tca gtc att gtc atg ctg gtt tct 3661
 Leu Met Val His Glu Gln Lys Val Ser Val Ile Val Met Leu Val Ser
 905 910 915

gag gct gag atg gag aag gtgagaagag gggtgggtgc ccccgaggca 3709
 Glu Ala Glu Met Glu Lys
 920

gtgtggggtg gcagggcagg ggatcctgga aaaccaggtc tgtcttggt tatctgtccc 3769

tcag caa aaa gtg gca cgc tac ttc ccc acc gag agg ggc cag ccc atg 3818
 Gln Lys Val Ala Arg Tyr Phe Pro Thr Glu Arg Gly Gln Pro Met
 925 930 935

gtg cac ggt gcc ctg agc ctg gca ttg agc agc gtc cgc agc acc gaa 3866
 Val His Gly Ala Leu Ser Leu Ala Leu Ser Ser Val Arg Ser Thr Glu
 940 945 950 955

acc cat gtg gag cgc gtg ctg agc ctg cag ttc cga gac cag agc ctc 3914
 Thr His Val Glu Arg Val Leu Ser Leu Gln Phe Arg Asp Gln Ser Leu
 960 965 970

aag cgc tct ctt gtg cac ctg cac ttc ccc act tgg cct gag tt 3958
Lys Arg Ser Leu Val His Leu His Phe Pro Thr Trp Pro Glu Leu
975 980 985

gtgagtccac tgctctggat ggtggttggg ggtctaagtg ctgtccagtc cttggtgctg 4018

ggagggatga gaggctcagg tcaggcctgg ctcataggct cttcctggcc ccacccctgctc 4078

ccacag a ggc ctg ccc gac agc ccc agc aac ttg ctg cgc ttc atc cag 4127
Gly Leu Pro Asp Ser Pro Ser Asn Leu Leu Arg Phe Ile Gln
990 995 1000

gag gtg cac gca cat tac ctg cat cag cgg ccg ctg cac acg ccc atc 4175
Glu Val His Ala His Tyr Leu His Gln Arg Pro Leu His Thr Pro Ile
1005 1010 1015

att gtg cac tgc ag gtagagggtg ggcctgaggg tctctcctct atgggctctt 4229
Ile Val His Cys Ser
1020

ggcctagcct cataccccgg cctcataacc ccttcttgge acag c tct ggt gtg 4283
Ser Gly Val

ggc cgc acg gga gcc ttt gca ctg ctc tat gca gct gtg cag gag gtg 4331
Gly Arg Thr Gly Ala Phe Ala Leu Leu Tyr Ala Ala Val Gln Glu Val
1025 1030 1035 1040

gag gct ggg aac gga atc cct gag ctg cct cag ctg gtg cgg cgc atg 4379
Glu Ala Gly Asn Gly Ile Pro Glu Leu Pro Gln Leu Val Arg Arg Met
1045 1050 1055

cgg cag cag aga aag cac atg ctg cag gag aag gtgatgatct gggcatatgg 4432
Arg Gln Gln Arg Lys His Met Leu Gln Glu Lys
1060 1065

ggctgggatg ggccttctgt cccagggatga cgggccctg cccagctgac ctggccaaat 4492

gcacctgtgc ag ctg cac ctc agg ttc tgc tat gag gca gtg gtg aga cac 4543
Leu His Leu Arg Phe Cys Tyr Glu Ala Val Val Arg His
1070 1075 1080

gtg gag cag gtc ctg cag cgc cat ggt gtg cct cct cca tgc aaa ccc	4591
Val Glu Gln Val Leu Gln Arg His Gly Val Pro Pro Pro Cys Lys Pro	
1085 1090 1095	
ttg gcc agt gca agc atc agc cag aag gtgaggaagg ttccgtggaa	4638
Leu Ala Ser Ala Ser Ile Ser Gln Lys	
1100 1105	
gctgctggga gagccacagc cttgggaatc cctctcctca ctcactctgt cttctcag	4696
aac cac ctt cct cag gac tcc cag gac ctg gtc ctc ggt ggg gat gtg	4744
Asn His Leu Pro Gln Asp Ser Gln Asp Leu Val Leu Gly Gly Asp Val	
1110 1115 1120	
ccc atc agc tcc atc cag gcc acc att gcc aag ctc agc att cgg cct	4792
Pro Ile Ser Ser Ile Gln Ala Thr Ile Ala Lys Leu Ser Ile Arg Pro	
1125 1130 1135	
cct ggg ggg ttg gag tcc ccg gtt gcc agc ttg cca ggc cct gca gag	4840
Pro Gly Gly Leu Glu Ser Pro Val Ala Ser Leu Pro Gly Pro Ala Glu	
1140 1145 1150	
ccc cca ggc ctc ccg cca gcc agc ctc cca gag tct acc cca atc cca	4888
Pro Pro Gly Leu Pro Pro Ala Ser Leu Pro Glu Ser Thr Pro Ile Pro	
1155 1160 1165	
tct tcc tcc cca ccc ccc ctt tcc tcc cca cta cct gag gct ccc cag	4936
Ser Ser Ser Pro Pro Pro Leu Ser Ser Pro Leu Pro Glu Ala Pro Gln	
1170 1175 1180 1185	
cct aag gag gag ccg cca gtg cct gaa gcc ccc agc tcg ggg ccc ccc	4984
Pro Lys Glu Glu Pro Pro Val Pro Glu Ala Pro Ser Ser Gly Pro Pro	
1190 1195 1200	
tcc tcc tcc ctg gaa ttg ctg gcc tcc ttg acc cca gag gcc ttc tcc	5032
Ser Ser Ser Leu Glu Leu Leu Ala Ser Leu Thr Pro Glu Ala Phe Ser	
1205 1210 1215	
ctg gac agc tcc ctg cgg ggc aaa cag cgg atg agc aag cat aac ttt	5080
Leu Asp Ser Ser Leu Arg Gly Lys Gln Arg Met Ser Lys His Asn Phe	
1220 1225 1230	

ctg cag gcc cat aac ggg caa ggg ctg cgg gcc acc cgg cct tct gac 5128
 Leu Gln Ala His Asn Gly Gln Gly Leu Arg Ala Thr Arg Pro Ser Asp
 1235 1240 1245

gac ccc ctc agc ctt ctg gat cca ctc tgg aca ctc aac aag acc 5173
 Asp Pro Leu Ser Leu Leu Asp Pro Leu Trp Thr Leu Asn Lys Thr
 1250 1255 1260

tgaacaggtt ttgcctacct ggtccttaca ctacatcatc atcatctcat gccacactgc 5233

ccacaccag cagagcttct cagtgggcac agtctcttac tcccatttct gctgcctttg 5293

gccctgcctg gccacgctg caccctgtg ggggtgaaat gtactgcagg ctctgggtca 5353

ggttctgctc ctttatggga cccgacattt ttcagctctt tgctattgaa ataataaacc 5413

accctgttct gtgaaaaaaaa aaa 5436

<210> 6

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> sense primer for amplification of HD-PTP gene

<400> 6

ctacagccac atcccccccc a

21

<210> 7

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> antisense primer for amplification of HD-PTP gene

<400> 7

cctcactctc ctccctgcgg

20

<210> 8
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> sense primer for amplification of HD-PTP gene position
1-248 containing exon 1.

<400> 8
gtggagggtg gcacagaggg

20

<210> 9
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> antisense primer for amplification of HD-PTP gene position
1-248 containing exon 1.

<400> 9
gtttggaatg gacatttcac tcct

24

<210> 10
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> sense primer for amplification of HD-PTP gene position
468-730 containing exon 2.

<400> 10
ctttgacatg gtcagagttg gat

23

<210> 11

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> antisense primer for amplification of HD-PTP gene position
468-730 containing exon 2.

<400> 11

cacagtagat gacctccaca ct

22

<210> 12

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> sense primer for amplification of HD-PTP gene position
934-1311 containing exon 3.

<400> 12

ggctgcctgt acaatccaca ac

22

<210> 13

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> antisense primer for amplification of HD-PTP gene position
934-1311 containing exon 3.

<400> 13

cccagctcga gtctcccaat g

21

<210> 14

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> sense primer for amplification of HD-PTP gene position
1323-1782 containing exon 4 and exon 5.

<400> 14

cctcaccgac cactgctgcc

20

<210> 15

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> antisense primer for amplification of HD-PTP gene position
1323-1782 containing exon 4 and exon 5.

<400> 15

cgggaaagga cagaggactg g

21

<210> 16

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> sense primer for amplification of HD-PTP gene position
1808-2293 containing exon 6 and part of exon 7.

<400> 16

gttcacctgg agctggccct t

21

<210> 17

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> antisense primer for amplification of HD-PTP gene position

1808-2293 containing exon 6 and part of exon 7.

<400> 17

gtggagcggg gtggcacttc

20

<210> 18

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> sense primer for amplification of HD-PTP gene position
2252-2590 containing part of exon 7.

<400> 18

cgactgcctg acaccttcct g

21

<210> 19

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> antisense primer for amplification of HD-PTP gene position
2252-2590 containing part of exon 7.

<400> 19

tgagaattga ggaggtggg c

21

<210> 20

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> sense primer for amplification of HD-PTP gene position
2500-2888 containing part of exon 7.

<400> 20

acagcatggc gtggtgagca g 21

<210> 21

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> antisense primer for amplification of HD-PTP gene position
2500-2888 containing part of exon 7.

<400> 21

gccccaaacgc ttgtgaagga tg 22

<210> 22

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> sense primer for amplification of HD-PTP gene position
2787-3217 containing part of exon 7.

<400> 22

agcaccactt ctcttctggg at 22

<210> 23

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> antisense primer for amplification of HD-PTP gene position
2787-3217 containing part of exon 7.

<400> 23

ggactggcca gcagacgagg 20

<210> 24

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> sense primer for amplification of HD-PTP gene position
3175-3599 containing part of exon 7.

<400> 24

agccccctctt accattcgag g

21

<210> 25

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> antisense primer for amplification of HD-PTP gene position
3175-3599 containing part of exon 7.

<400> 25

catcatgttc ctgcgcatct tg

22

<210> 26

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> sense primer for amplification of HD-PTP gene position
3553-3948 containing part of exon 7.

<400> 26

tctggacact gtctggcgag a

21

<210> 27

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> antisense primer for amplification of HD-PTP gene position
3553-3948 containing part of exon 7.

<400> 27

caagacagac ctggttttcc ag

22

<210> 28

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> sense primer for amplification of HD-PTP gene position
3937-4179 containing exon 8.

<400> 28

aggtctgtct tggcttatct gt

22

<210> 29

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> antisense primer for amplification of HD-PTP gene position
3937-4179 containing exon 8.

<400> 29

ccaaccacca tccagagcag t

21

<210> 30

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> sense primer for amplification of HD-PTP gene position
4206-4440 containing exon 9.

<400> 30

tgctgggagg gatgagagcc

20

<210> 31

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> antisense primer for amplification of HD-PTP gene position
4206-4440 containing exon 9.

<400> 31

cgggtatga ggctaggcca

20

<210> 32

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> sense primer for amplification of HD-PTP gene position
4400-4688 containing exon 10.

<400> 32

ggtctctect ctatgggctc t

21

<210> 33

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> antisense primer for amplification of HD-PTP gene position
4400-4688 containing exon 10.

<400> 33
gtgcatttgg ccaggtcagc t 21

<210> 34
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> sense primer for amplification of HD-PTP gene position
4626-4878 containing exon 11.

<400> 34
gctgggatgg gccttctgtc 20

<210> 35
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> antisense primer for amplification of HD-PTP gene position
4626-4878 containing exon 11.

<400> 35
agagtgagtg aggagaggga tt 22

<210> 36
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> sense primer for amplification of HD-PTP gene position
4801-5214 containing part of exon 12.

<400> 36
cagccagaag gtgaggaagg tt 22

<210> 37

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> antisense primer for amplification of HD-PTP gene position
4801-5214 containing part of exon 12.

<400> 37

tctggggtca aggaggccag c

21

<210> 38

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA for the multicloning site (XhoI NotI XbaI KpnI BamHI) linker

<400> 38

tcgagcggcc gctctagagg taccg

25

<210> 39

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA for the multicloning site (XhoI NotI XbaI KpnI BamHI) linker

<400> 39

gatccggtac ctctagagcg gccgc

25

<210> 40

<211> 297

<212> DNA

<213> Homo sapiens

<220>

<221> exon

<222> (1)..(131)

<220>

<221> intron

<222> (132)..(297)

<400> 40

tgcagcagct gcgggagtgg ccgggtggcc ggcggtgcc acccgcc atg gag gcc 56
Met Glu Ala

1

gtg ccc cgc atg ccc gtg atc tgg ctg gac ctg aag gag gcc ggt gac 104
Val Pro Arg Met Pro Val Ile Trp Leu Asp Leu Lys Glu Ala Gly Asp
5 10 15

ttt cac ttc cag cca gct gtg aag aag gtgagcttgc cttccatctt 151
Phe His Phe Gln Pro Ala Val Lys Lys
20 25

ccccctatc cgccgcgtat ctccgtcgtc tgcaagcgcg tgacgcccc ctgcgcgccc 211

atcacctcct caggcccggt cgcagggtcc ggtgaggcca ggaggggcct tcgccggttt 271

tcctcagcct gtccacacac ccggcg 297

<210> 41

<211> 1024

<212> DNA

<213> Homo sapiens

<220>

<221> exon

<222> (404)..(478)

<220>

<221> intron

<222> (1)..(403)

<220>

<221> intron

<222> (479)..(1024)

<400> 41

tcttgctgc agttgctatc agcttaggtt atggtagtca cccacaacg cagcaggccc 60
aatcttctag ggaggcttga gttcgtgacc ctttctggaa tcccttatta ggacccatt 120
tcagtccgct gagttccttg gcttctgag ctggcagact cctggacttt gtagagagac 180
tgcagcgta ggaggacctg ccagagtgtt ggtgcagggt caaggctctc gtcagcattt 240
tggtgctgcc tgttcaggag gcgaagtcc tggtgtcca ggagacacgc ttatggttgt 300
aggtctgcac ttaatcctga gtcttggttg gtgcttgccc aggactcaga gggcaaggcg 360
gggtcttct ctgccactgg cttatgttct tctctctctg tag ttt gtc ctg aag 415
Phe Val Leu Lys
1

aat tat gga gag aac cca gaa gcc tac aat gaa gaa ctg aag aag ctg 463
Asn Tyr Gly Glu Asn Pro Glu Ala Tyr Asn Glu Glu Leu Lys Lys Leu
5 10 15 20

gag ttg ctc aga cag gtaggaggat agtattatct tttatgcat gggtagacag 518
Glu Leu Leu Arg Gln
25

gattggtttg ataggagat aaagaaactg cctaggctgg gcatggtggc tcaacgccta 578

taatcccacc actttgggag gccgaggtgg gcagatcatt tgaggtcagg agtttgagac 638
cagcctggcc aacatggtga aaccccatct ctactaaaaa tacaaaaatt aggcaggtgt 698
ggtgtcatgt gcctgtagtt ccggctactc aggagtctga agcaggataa ttgcttcaac 758
ccaggaggtg gaggttgcatt tgagctgaga tcatgccact gcactccagc ctgggtgata 818
gagcgagact ccatctcaaa aacaacaaaa aagaaaccac ctggcccttt ctagcttttg 878
atgtaazgtg aaagacagct ggatgtgtga ctcatgccta tgatcccacc actttgggag 938
gccaaaggcag gaggattggt tcctcctgga gttcaggccc aggagtttga gacccctca 998
acattgtgag gccctgtctc tataga 1024

国際調査報告

国際出願番号 PCT/JP00/02455

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl⁷ C12N15/55, 9/16, C07K16/40, C12Q1/68,
A61K38/46

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl⁷ C12N15/00-15/90

最小限資料以外の資料で調査を行った分野に含まれるもの

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

GENBANK/EMBL/DDBJ/GENESEQ
PIR/SWISSPROT/GENESEQ

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X	The Journal of Biological Chemistry, 273(33), Aug. 1998 Linguang Cao et al., "A Novel Putative Protein-tyrosine Phosphatase contains a BR01-like domain and suppresses Ha-ras-mediated transformation", p. 21077-21083	1-13, 15-18, 20
X	WO, 98/49317, A2 (SUGEN INC.) 05. 11月. 1998 (05. 11. 98) p. 160-164 &EP, 979288, A2 &AU, 9872600, A	1-13, 15-18, 20

☐ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日

21. 06. 00

国際調査報告の発送日

04.07.00

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

郵便番号 100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

鈴木 恵理子

印

4B

9838

電話番号 03-3581-1101 内線 3448

法第8条第3項（PCT17条(2)(a)）の規定により、この国際調査報告は次の理由により請求の範囲の一部について作成しなかった。

- 請求の範囲14、19は、人の診断方法であるから、PCT17条(2)(a)(I)及びPCT規則39.1(IV)の規定により、この国際調査機関が調査することを要しない対象に係るものである。

3. ☐ 請求の範囲 _____ は、従属請求の範囲であって P C T 規則 6.4 (a) の第 2 文及び第 3 文の規定に従って記載されていない。

次に述べるようにこの国際出願に二以上の発明があるとの国際調査機関は認めた。

1. ☐ 出願人が必要な追加調査手数料をすべて期間内に納付したので、この国際調査報告は、すべての調査可能な請求の範囲について作成した。
2. ☐ 追加調査手数料を要求するまでもなく、すべての調査可能な請求の範囲について調査することができたので、追加調査手数料の納付を求めなかった。
3. ☐ 出願人が必要な追加調査手数料を一部のみしか期間内に納付しなかったので、この国際調査報告は、手数料の納付のあった次の請求の範囲のみについて作成した。
4. ☐ 出願人が必要な追加調査手数料を期間内に納付しなかったので、この国際調査報告は、請求の範囲の最初に記載されている発明に係る次の請求の範囲について作成した。

☐ 追加調査手数料の納付と共に出願人から異議申立てがあった。

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP00/02455

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl.⁷ C12N15/55, 9/16, C07K16/40, C12Q1/68,
A61K38/46

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl.⁷ C12N15/00-15/20

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

GENBANK/EMBL/DDBJ/GENESFQ
PIR/SWISSPROT/GENESEQ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	The Journal of Biological Chemistry, 273(33), Aug. 1998 Linguang Cao et al., "A Novel Putative Protein-tyrosine Phosphatase contains a BRO1-like domain and suppresses Ha-ras-mediated transformation", p.21077-21083	1-13, 15-18, 20
X	WO, 98/49317, A2 (SUGEN INC.), 05 November, 1998 (05.11.98), p.160-164 & EP, 979288, A2 & AU, 9872600, A	1-13, 15-18, 20

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not
considered to be of particular relevance

"E" earlier document but published on or after the international filing
date

"L" document which may throw doubts on priority claim(s) or which is
cited to establish the publication date of another citation or other
special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other
means

"P" document published prior to the international filing date but later
than the priority date claimed

"T" later document published after the international filing date or
priority date and not in conflict with the application but cited to
understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be
considered novel or cannot be considered to involve an inventive
step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be
considered to involve an inventive step when the document is
combined with one or more other such documents, such
combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
21 June, 2000 (21.06.00)

Date of mailing of the international search report
04 July, 2000 (04.07.00)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP00/02455

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 14,19
because they relate to subject matter not required to be searched by this Authority, namely:
The subject matter of claims 14,19 relates to a method for diagnosis of the human body, which does not require an international search report by this International Search Authority in accordance with PCT Article 17(2)(a)(i) and Rule 39.1(iv) ..
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.